

AFIT/GE/ENG/03-18



REDUNDANT DISCRETE WAVELET TRANSFORM BASED
SUPER-RESOLUTION USING SUB-PIXEL IMAGE REGISTRATION

THESIS
Daniel L. Ward
Second Lieutenant, USAF

AFIT/GE/ENG/03-18

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

AFIT/GE/ENG/03-18

REDUNDANT DISCRETE WAVELET TRANSFORM BASED
SUPER-RESOLUTION USING SUB-PIXEL IMAGE
REGISTRATION

THESIS

Presented to the Faculty
Graduate School of Engineering and Management
Air Force Institute of Technology
Air University
Air Education and Training Command
In Partial Fulfillment of the Requirements for the
Degree of Master of Science in Electrical Engineering

Daniel L. Ward, B.S.E.E.
Second Lieutenant, USAF

March, 2003

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

REDUNDANT DISCRETE WAVELET TRANSFORM BASED
SUPER-RESOLUTION USING SUB-PIXEL IMAGE
REGISTRATION

THESIS

Daniel Lee Ward, B.S.E.E.

Second Lieutenant, USAF

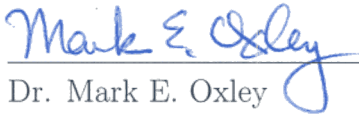
Approved:



Maj Roger L. Claypoole, Jr., Ph.D.
Thesis Advisor

6 MAR 03

Date



Dr. Mark E. Oxley
Committee Member

6 Mar 03

Date



Maj Matthew E. Goda, Ph.D.
Committee Member

6 MAR 03

Date

Acknowledgements

I would like to thank my committee for helping me to complete this document. Many thanks to Major Claypoole for focusing my ideas, and having confidence in my work. Thanks to the VIP lab crew, and the other AFIT students who frequented the lab.

Daniel L. Ward

Table of Contents

	Page
Acknowledgements	iv
List of Figures	vii
List of Tables	ix
Abstract	x
 I. Introduction	 1-1
1.1 Problem Statement	1-1
1.2 Scope	1-2
1.3 Thesis Organization	1-2
 II. Background	 2-1
2.1 The Discrete Wavelet Transform	2-1
2.1.1 The Multiresolution Expansion Set	2-1
2.1.2 Discrete Wavelet Transform Implementation	2-3
2.1.3 Two-Dimensional Discrete Wavelet Transform	2-5
2.1.4 Properties of the Wavelet Transform	2-6
2.1.5 Orthogonal vs. Biorthogonal Wavelet Filters	2-9
2.2 Super-Resolution	2-10
2.2.1 Low-Resolution Frame Model	2-11
2.2.2 Frame Alignment and Pixel Grid Conversion	2-12
2.2.3 Super-Resolution Estimation	2-16
2.2.4 Summary	2-17
 III. Methodology	 3-1
3.1 The Test Data - Half Pixel Shifts	3-1
3.2 The Wavelet Super-Resolution Algorithm	3-2
3.2.1 Pre-Wavelet Domain Operations	3-2
3.2.2 Wavelet Domain Operations	3-7
3.2.3 Post Wavelet Domain Processing	3-12
3.3 Summary	3-15
 IV. Results	 4-1
4.1 The Test Images	4-1
4.2 The Wavelets	4-1
4.3 The Comparisons	4-2
4.4 Analysis	4-3
4.5 Summary	4-16

	Page
V. Discussion and Future Work	5-1
5.1 Contributions Of This Thesis	5-1
5.2 Potential For Future Research	5-2
Appendix A. Determination of Missing Pixel Locations	A-1
Bibliography	BIB-1

List of Figures

Figure		Page
2.1.	Wavelet Subspaces	2-2
2.2.	Wavelet Subspace Iterations	2-3
2.3.	Analysis Filter Bank	2-4
2.4.	Redundant DWT Analysis Filter Bank	2-4
2.5.	Synthesis Filter Bank	2-5
2.6.	Two Dimensional Discrete Wavelet Transform	2-6
2.7.	Discrete Wavelet Transform Applied to an Image	2-7
2.8.	Low High Subband at Different Scales of RDWT	2-8
2.9.	Orthogonal and Biorthogonal Wavelet Filters	2-10
2.10.	Band-Limited Finite CCD Array	2-12
2.11.	Overlapping Frames in a Video Sequence	2-13
2.12.	Low Resolution Pixel Alignment	2-14
2.13.	Sub-Pixel Shift	2-15
2.14.	High-Resolution Pixel Alignment Due To Sub-Pixel Shift	2-15
2.15.	Mendoza Super-Resolution Algorithm	2-16
3.1.	Half Pixel Shift and Quincunx Sampling	3-2
3.2.	<i>Traditional vs. My Super-Resolution Pixel Association</i>	3-3
3.3.	Half Pixel Shift Application	3-3
3.4.	Combining Test Frames Into High Resolution Grid	3-5
3.5.	Pre-Wavelet Domain Operations	3-8
3.6.	Missing Coefficient and the Four Nearest Neighbors	3-9
3.7.	Pixel Groupings	3-10
3.8.	Re-Rotation of High-Resolution Image	3-12
3.9.	Pixel Correlations	3-13
3.10.	Erroneous Edges in HH Subband	3-14
3.11.	Correcting the HH Subband	3-14
4.1.	Original and Super-Resolution Test Images	4-6
4.2.	Original and Super-Resolution Test Images	4-7
4.3.	Original and Super-Resolution Test Images	4-8
4.4.	Circles Comparison	4-9
4.5.	Rectangles Comparison	4-10
4.6.	Pinwheel Comparison	4-11
4.7.	Lenna Comparison	4-12
4.8.	Cameraman Comparison	4-13

Figure		Page
4.9.	Aerial Comparison	4-14
4.10.	Interpolation Comparison with Circles Image	4-15
4.11.	Interpolation Comparison with Rectangle Image	4-15
4.12.	Interpolation Comparison with Lenna Image	4-16
4.13.	Interpolation Comparison with Cameraman Image	4-17
4.14.	Pinwheel Before and After Post-Processing	4-17
1.1.	Convolution of Mask with Sub-Band	A-2

List of Tables

Table		Page
4.1.	Wavelets Used for Testing.	4-2
4.2.	Mean Square Error Comparison of Wavelet Filters.	4-16
4.3.	Mean Square Error Comparison of the Daubechies 7,9 Wavelet with Traditional Interpolation Methods.	4-18
4.4.	L-Infinity Comparison of Wavelet Filters	4-18
4.5.	L-Infinity Comparison of the Daubechies 7,9 Wavelet with Tra- ditional Interpolation Methods.	4-18
4.6.	Mean Square Error Comparison of Post-Processing.	4-18
4.7.	L-Infinity Comparison of Post-Processing	4-18

Abstract

The limited resolution of video imagery taken by aircraft, over geographical areas of interest, hinders the accurate extraction of useful information. The frame resolution of the video is determined by the camera that created it. Information exists about the camera which can be used to increase frame resolution beyond the resolution capability of the camera. This is achieved by a process called *super-resolution*, which uses multiple low-resolution video frames to create one high-resolution image.

Paramount to this super-resolution process is the alignment of the low resolution frames. Frame alignment is achieved by image registration. Many good image registration algorithms utilize a *wavelet domain* analysis of the images. It is theorized that an accurate super-resolution algorithm can be created that also utilizes wavelet domain analysis. By achieving both image registration and super-resolution in the wavelet domain, a more computationally efficient overall system may be developed.

This thesis explores the possibility of wavelet-based super-resolution. An algorithm is developed that restores resolution lost by the CCD array of the camera. Two low-resolution frames, registered at a specific shift from one another, are used to create one high-resolution image. The algorithm is compared to other traditional interpolation techniques, and different wavelets systems are applied to determine which one works best.

REDUNDANT DISCRETE WAVELET TRANSFORM BASED SUPER-RESOLUTION USING SUB-PIXEL IMAGE REGISTRATION

I. Introduction

1.1 Problem Statement

Video imagery taken by aircraft, over geographical areas of interest, provides the warfighter with the most up-to-date information. The data can be used to determine potential targets, to locate and determine the number of vehicles and/or buildings in an area, as well as a variety of other information. The accuracy of this information depends upon the resolution of the video imagery. Higher resolution typically means more accurate information.

The camera, which provides the video imagery, has a limited resolution. Each frame in the sequence is taken at a resolution determined by the camera's optics. However, information exists about the camera, which can be used to increase frame resolution beyond the limitations of the camera. This is achieved through the accurate alignment of adjacent frames in the video sequence. The process of using multiple aligned frames to create one high resolution image is called *super-resolution*.

Existing super-resolution techniques incorporate specific knowledge about the camera's optics, and CCD array density. They determine how the low resolution image was created, by estimating the distortion operations, such as blurring and warping, performed by the camera. Most of these algorithms require the accurate alignment of multiple frames. Frame alignment is achieved by image registration. Many good registration algorithms use wavelets. This thesis develops a super-resolution algorithm that also uses wavelets. By achieving registration and super-resolution using wavelets, a more computationally efficient overall system may be developed.

1.2 *Scope*

In a video sequence, where a camera is moving, each frame has a position over an area relative to the frames adjacent to it. The process of determining the position of one frame relative to the other is called image registration. Using image registration, the super-resolution algorithm developed by this thesis places two adjacent frames together into a high resolution grid. The algorithm assumes that the frames are at specific positions relative to one another. Once the frames are placed in the high resolution grid, the algorithm use the redundant discrete wavelet transform to interpolate the remaining values in the grid.

Different wavelets are experimented with to determine which ones are best at interpolation. The wavelet interpolation technique is then compared to traditional interpolation techniques.

1.3 *Thesis Organization*

Chapter 2 discusses the unique properties of the wavelet transform. The discrete wavelet transform and redundant discrete wavelet transform are described in detail. The difference between orthogonal and biorthogonal wavelets is presented next. Following this is a discussion on classic super-resolution algorithms.

Chapter 3 breaks down the super-resolution algorithm step by step. It begins with an explanation of sub-pixel image registration and the creation of the test data. This is followed by a discussion on how and why the test data is pieced together into the high resolution grid. Next, the wavelet interpolation process is described. Finally, a unique post-processing technique is developed and explained. The chapter concludes with a summary of the algorithm's capability.

Chapter 4 presents the results of the super-resolution algorithm, and analyzes its performance. It compares restored super-resolution images, generated by the algorithm, to the original high resolution images and test data. The performance of a number of different wavelets is analyzed to determine which wavelet performs

best. This is done qualitatively by showing specific portions of certain images, and quantitatively by computing the mean square error and L-infinity error for each wavelet image. The wavelet that performs the best is then compared to other traditional interpolation techniques, including cubic spline, nearest neighbor, and linear interpolation.

Chapter 5 discusses research contributions and the potential for future work.

II. Background

Since the general introduction of wavelets by Ingrid Daubechies in 1988 [6], wavelets have become an essential part of image and video processing. Applications include image registration, compression, and noise reduction. The algorithm developed in this thesis applies wavelets to the problem of super-resolution. Therefore this chapter discusses both the basics of the wavelet transform and existing super-resolution techniques.

2.1 The Discrete Wavelet Transform

The wavelet transform expands a signal, or function, into the wavelet domain. As with any transform, like the Fourier or Gabor Transforms, the goal of expanding a signal is to obtain information that is not apparent, or can not be deduced, from the signal in its original domain (usually space, time or distance). The wavelet transform is used in this thesis because of the unique information it provides about images. The algorithm developed in Chapter 3 is much more effective when processing images in the wavelet domain. However, before this can be discussed, it is first necessary to understand how the wavelet transform expands a signal into the wavelet domain.

2.1.1 The Multiresolution Expansion Set. An expansion set is a set of functions, $\varphi_k(t), k \in \mathcal{Z}$, that can represent any signal in a vector space, V , as

$$f(t) = \sum_k a_k \varphi_k(t), \quad k \in \mathcal{Z}, \quad f(t) \in V. \quad (2.1)$$

$\varphi_k(t)$ spans the vector space V , and is said to form a *basis set* if the a_k are unique for any given $f(t)$ [4].

The wavelet expansion set includes a set of two functions, the low-pass scaling function, $\phi(t)$, the band-pass wavelet function, $\psi(t)$, and their translations. These functions form the basis set, and divide the original subspace, V , into two different

subspaces, V_1 and W_1 , where $V_1 \perp W_1$ and, $V_1 \oplus W_1 \subset V$ (See Figure 2.1). Instead of

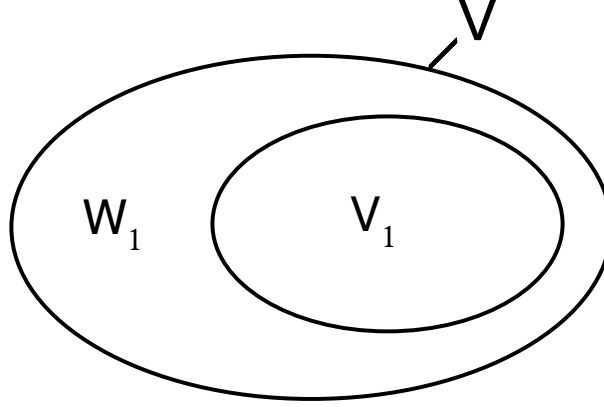


Figure 2.1: Wavelet Subspaces. The $\phi(t)$ and $\psi(t)$ Functions project a signal into subspaces V_ϕ and W_ψ respectively.

representing the signal with one set of coefficients, the signal is now represented by two, c_k and d_k . The subscript k represents each translation of the signal. Equations 2.2 and 2.3 show this new dual expansion set:

$$f(t) = \sum_k c_k \phi_k(t) + \sum_k d_k \psi_k(t) \quad (2.2)$$

or

$$f(t) = \sum_k c(k) \phi(t - k) + \sum_k d(k) \psi(t - k). \quad (2.3)$$

The wavelet transform can expand the signal into multiple scales. This is done by *dilating* the scaling and wavelet functions according to the appropriate scale. Derivation of the dilation is not pertinent to this thesis, but is shown in Equations 2.4 through 2.7. Dilation of the functions is represented by a subscript m , or number for the appropriate scale, as seen in Equations 2.4 through 2.7. The coefficients c_k are then expanded in the same way as the original signal.

$$\text{Scale One : } f(t) = \sum_k c_{1,k} \phi_{1,k}(t) + \sum_k d_{1,k} \psi_{k,1}(t) \quad (2.4)$$

$$\text{Scale Two : } c_{1,k} = \sum_k c_{2,k} \phi_{2,k}(t) + \sum_k d_{2,k} \psi_{2,k}(t) \quad (2.5)$$

This process can be iterated to the desired number of scales.

$$\text{Scale } m : \quad c_{m,k} = \sum_k c_{m+1,k} \phi_{m+1,k}(t) + \sum_k d_{m+1,k} \psi_{m+1,k}(t) \quad (2.6)$$

$$f(t) \text{ Expansion} : \quad f(t) = \sum_k c_{m,k} \phi_{m,k}(t) + \sum_k \sum_{j=m}^{\infty} d_{j,k} \psi_{j,k}(t) \quad (2.7)$$

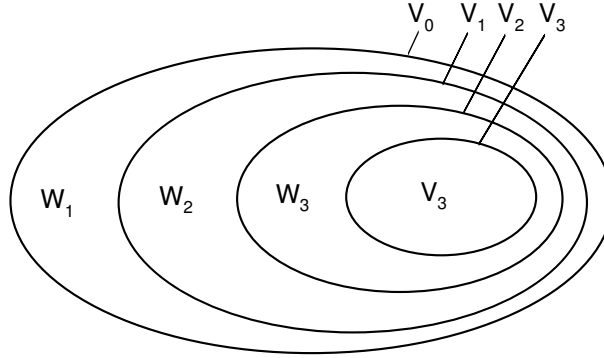


Figure 2.2: Wavelet Subspace Iterations. The $\phi(t)$ and $\psi(t)$ Functions project a signal into the desired number subspaces (scales).

Application of the wavelet transform to a signal is typically implemented with a filterbank. This is referred to as the *discrete wavelet transform* (DWT), and is discussed in the next section.

2.1.2 Discrete Wavelet Transform Implementation. The discrete wavelet transform uses low-pass and high-pass filters, $h(n)$ and $g(n)$, to expand a digital signal. They are referred to as *analysis filters*. These filters correspond to $\phi(t)$ and $\psi(t)$ of the previous section. The dilation performed for each scale is now achieved by a decimator. The coefficients c_k and d_k are produced by convolving the digital signal, with each filter, and then decimating the output. The c_k coefficients are produced by the low-pass filter, $h(n)$, and called *coarse coefficients*. The d_k coefficients are produced by the high-pass filter and called *detail coefficients*. Coarse coefficients provide information about low frequencies, and detail coefficients provide information about high frequencies. Coarse and detail coefficients are produced at

multiple scales by iterating the process on the coarse coefficients of each scale. The entire process is computed using a tree-structured filterbank, as seen in Figure 2.3.

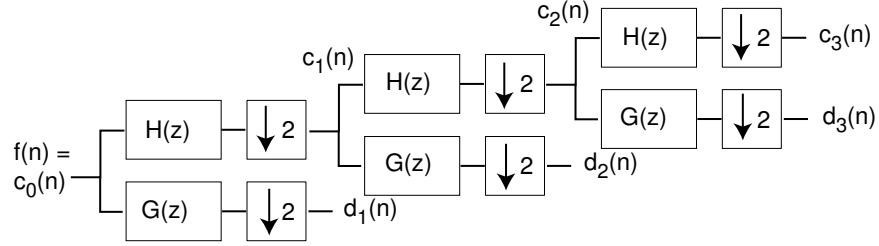


Figure 2.3: Analysis Filter Bank. The high and low pass filters divide the signal into a series of coarse and detail coefficients.

2.1.2.1 Redundant Discrete Wavelet Transform. The decimators remove redundant coefficients, which are not necessary to perfectly reconstruct the signal. This makes wavelet compression algorithms more computationally efficient. However, in some wavelet techniques the redundant coefficients are useful. For example, this thesis interpolates missing pixels of an image, in the wavelet domain, based on the values of the surrounding pixels. As much information as possible is needed to accurately interpolate the missing pixels. Decimation removes potentially valuable information. In cases like this it is beneficial to remove the decimators. This is known as the *redundant discrete wavelet transform* (RDWT), as shown in Figure 2.4.

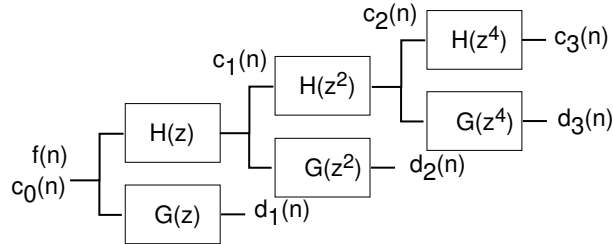


Figure 2.4: Redundant DWT Analysis Filter Bank. The decimators are removed so that more information about the signal is available.

2.1.2.2 Inverse Discrete Wavelet Transform. After analyzing, or processing, the signal in the wavelet domain it is often necessary to return the signal

back to its original domain. This is achieved using synthesis filters and expanders. The wavelet coefficients are applied to a synthesis filter bank to restore the original signal, as seen in Figure 2.5. In the case where the redundant discrete wavelet transform was used, the expanders in the synthesis filter bank are removed.

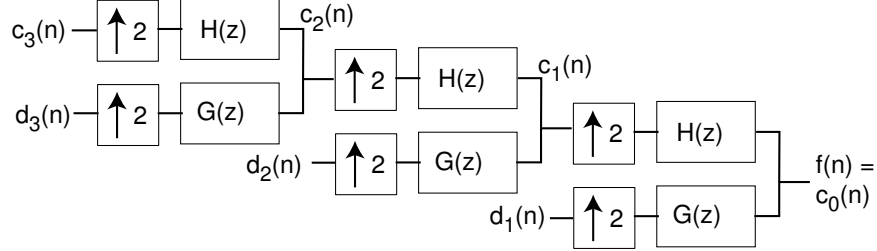


Figure 2.5: Synthesis Filter Bank. The high and low pass filters combine the coefficients into the original signal.

The discussion, up to now, has dealt exclusively with one-dimensional (1D) signals. This thesis deals exclusively with two-dimensional signals (2D), more specifically images. Therefore, it is necessary to extend the discussion into 2D signals.

2.1.3 Two-Dimensional Discrete Wavelet Transform. The two dimensional discrete wavelet transform is essentially a one dimensional analysis of a two dimensional signal. It only operates on one dimension at a time, by analyzing the rows and columns of an image in a separable fashion.

The first step applies the analysis filters to the rows of an image. This produces two new images, where one image is a set of coarse row coefficients, and the other a set of detail row coefficients. Next analysis filters are applied to the columns of each new image, to produce four different images called *subbands*. Rows and columns analyzed with a high pass filter are designated with an H. Likewise, rows and columns analyzed with a low pass filter are designated with an L. For example, if a subband image was produced using a high pass filter on the rows and a low pass filter on the columns, it is called the **HL** subband. Figure 2.6 shows this process in its entirety.

Each subband provides different information about the image. The LL subband is a coarse approximation of the image and removes all high frequency information.

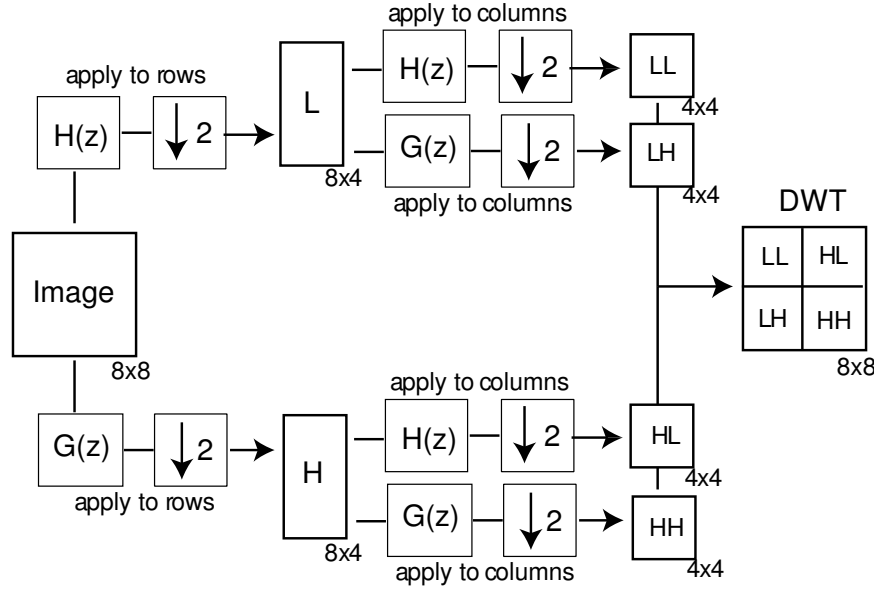


Figure 2.6: Two Dimensional Discrete Wavelet Transform. The high and low pass filters operate separably on the rows and columns to create four different subbands. An 8x8 image is used for example purposes only.

The LH subband removes high frequency information along the rows and emphasizes high frequency information along the columns. The result is an image in which vertical edges are emphasized. The HL subband emphasizes horizontal edges, and the HH subband emphasizes diagonal edges (See Figure 2.7). To compute the DWT of the image at the next scale the process is applied again to the LL subband. Figure 2.7 shows three iterations of this process on an image.

Just as in the 1D case, if the RDWT is desired for an image the decimators are removed. Each subband becomes the same size as the original image.

The wavelet domain representation of an image, or any signal, is useful for many applications, such as compression, noise reduction, image registration, and super-resolution. The properties which make it useful for super-resolution are explained in the next section.

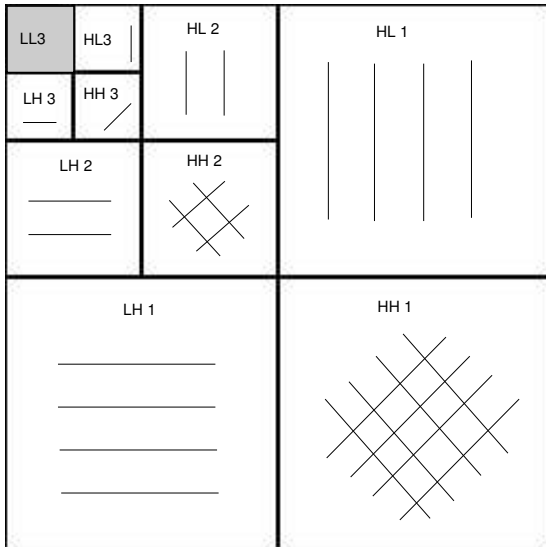
2.1.4 Properties of the Wavelet Transform. The wavelet domain representation of a signal is unique in that it simultaneously provides frequency and time (or



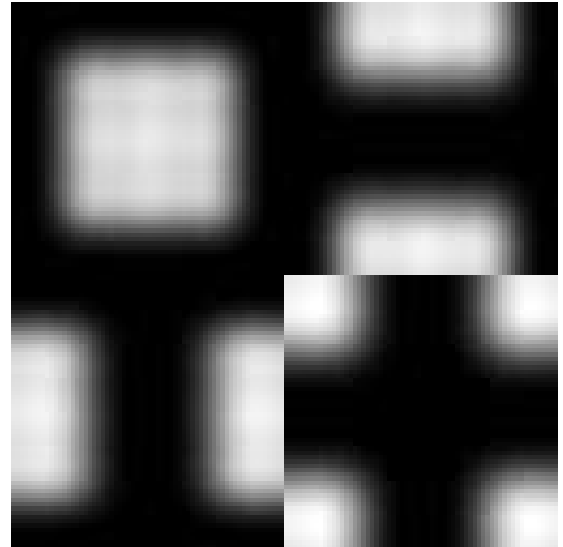
(a)



(b)



(c)



(d)

Figure 2.7: Discrete Wavelet Transform Applied to an Image. (a) The original image. (b) Three iterations of the wavelet transform are applied to the original image. The result shown here has coefficients at three different scales. (c) Each subband, except the LL subband, emphasizes edges in a certain direction. Vertical edges are emphasized in the LH subband, horizontal edges in the HL subband, and diagonal edges in the HH subband. (d) The frequency response of the each subband, shown here with only one iteration.

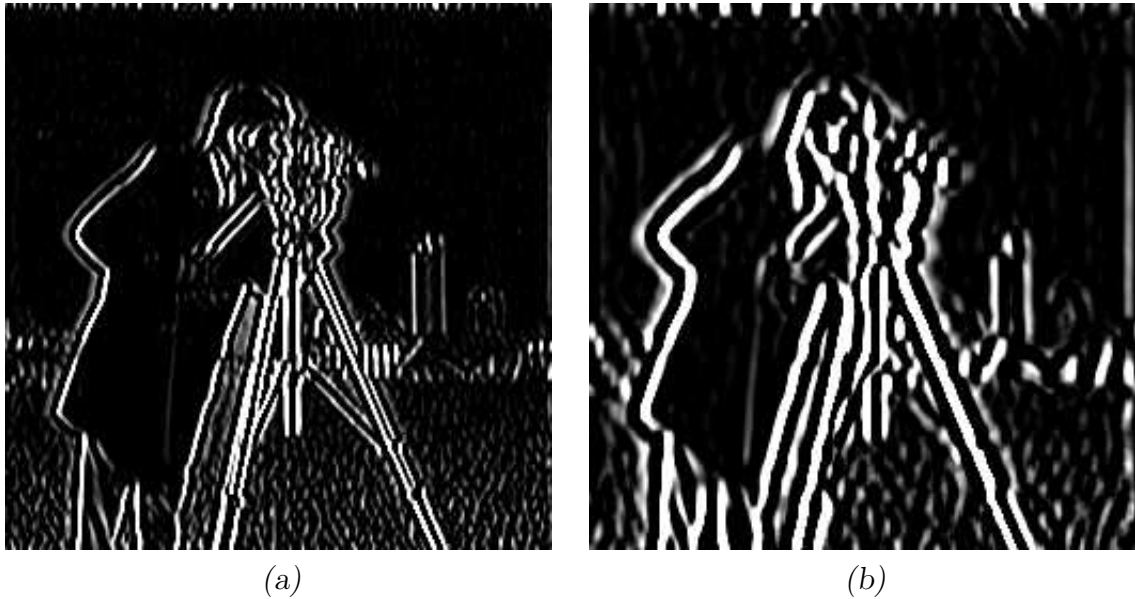


Figure 2.8: Low High Subband at Different Scales of RDWT. (a) LH Subband at scale 1. (b) LH Subband at scale 3. Each subband is the same size as the original image shown in Figure 2.7(a).

space) localization information. The frequency content of a signal is known, as well as when (or where) those frequencies occur. In an image, high frequency content exists along edges, and low frequency content exists where little or no edges occur. Therefore, each DWT subband provides information about where different types of edges do or do not exist. Coefficients in the subbands are large where edges do exist, and zero, or close to zero, where edges do not exist.

Since most real world images do not consist entirely of edges, most of the coefficients are nearly zero. This is referred to as a parsimonious representation [15]. Though parsimony is useful for such processes as compression, this thesis is more concerned with a property that derives itself from parsimony. Because small coefficients exist where edges are not present, and large coefficients only exist where they are present, coefficients of similar size tend to cluster together [14, 16]. The clustering of coefficients is extremely useful for the super-resolution of images, where pixel values are interpolated based upon adjacent pixel values. This helps to properly restore features in an image, such as edges, and is demonstrated in Chapter 4.

Certain properties of the wavelet transform depend upon which type of filters are used for analysis and synthesis. Filters used by the wavelet transform fall into two categories: orthogonal and biorthogonal. These categories are compared and contrasted in the next section.

2.1.5 Orthogonal vs. Biorthogonal Wavelet Filters. Classical wavelet systems use orthogonal filters, which require the filters to be orthogonal across both translation and scale. Taking the scaling and wavelet functions from Section 2.1.1, this means that

$$\langle \phi_{m,k}, \phi_{n,l} \rangle = \delta(k-l, m-n) \quad (2.8)$$

$$\langle \psi_{m,k}, \psi_{n,l} \rangle = \delta(k-l, m-n) \quad (2.9)$$

These requirements give a clean, robust system in which energy is preserved in accordance with Parseval's Theorem. However, the requirements also place large limitations on the possibilities of the system. The wavelet and scaling functions must have the same length and the length must be even [4]. These restrictions prevent linear phase analysis, except with the Haar wavelet. Linear phase analysis is crucial in image processing because the majority of the information in an image is contained in its phase [13]. Therefore, it is desirable, and in some cases necessary, to relax the restrictions of orthogonality. Biorthogonal filters are the result of this relaxation.

Biorthogonal filters are not required to be orthogonal across translation and scale. Instead of having each element in a basis set be orthogonal to each other, they are required to be orthogonal to the elements of a *dual basis*, $\{\tilde{\varphi}_{n,l}\}$.

$$\langle \phi_{m,k}, \tilde{\phi}_{n,l} \rangle = \delta(k-l, m-n) \quad (2.10)$$

$$\langle \psi_{m,k}, \tilde{\psi}_{n,l} \rangle = \delta(k-l, m-n) \quad (2.11)$$

The filters can be designed with different lengths, to include odd length filters. Perhaps most importantly, the filters can be designed symmetrically with linear phase. The only symmetric linear phase orthogonal filter is the Haar, which has a length of only two. Biorthogonal filters can be much larger than this. Larger filters generally correspond to smoother wavelet functions, and a more parsimonious signal representation [24]. Biorthogonal filters have been shown to give better space and

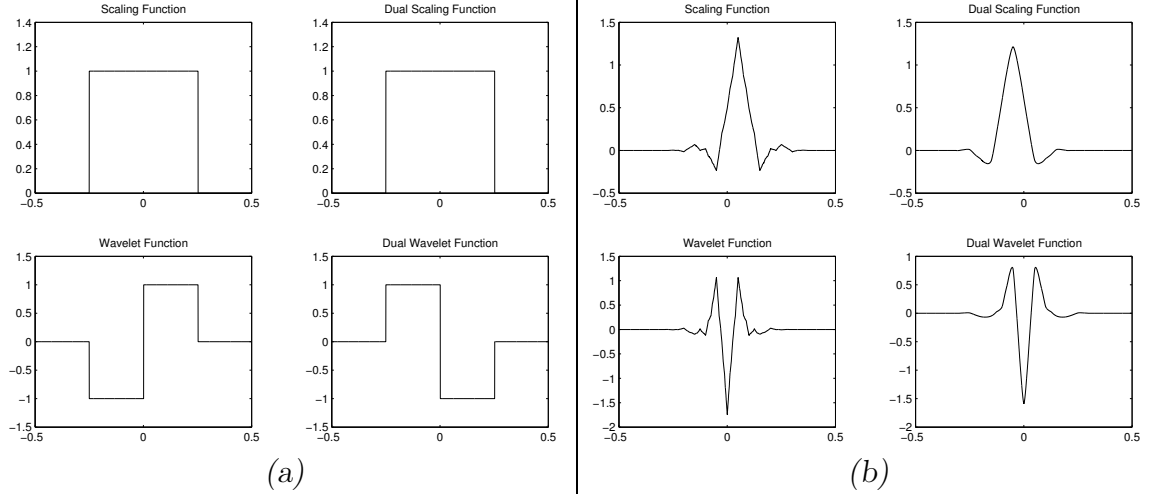


Figure 2.9: Orthogonal and Biorthogonal Wavelet Filters. (a) The orthogonal linear phase Haar filter. (b) The biorthogonal linear phase Daubechies 9,7 Filters.

spatial frequency localization [3], which is beneficial in both image registration and super-resolution. Therefore, the biorthogonal wavelet system is the system of choice for this thesis.

The wavelet transform has been explained in detail to provide the information necessary to understand the algorithm in Chapter 3. To better understand how the wavelet transform is useful for super-resolution, the process of super-resolution must now be discussed.

2.2 Super-Resolution

Super-Resolution is a process which creates a high-resolution image from several low-resolution images. The fundamental idea is to restore high frequency in-

formation that is lost by the image capturing process of the camera. This thesis specifically restores frequency loss to due aliasing of the camera's CCD array. Super-resolution was first suggested in 1955 [7]. For many years computing power severely limited the practical application of super-resolution. Since then, as computing power has increased, multiple algorithms have been proposed [1, 8–10, 12, 21, 22, 25], which incorporate techniques like Bayesian estimation, and Kalman filtering. A more extensive study of the history of super-resolution is available in [11, 23].

The majority of super-resolution algorithms depend upon relative motion, at the sub-pixel level, between the low-resolution images, though algorithms have been developed that do not [1]. Therefore, most super-resolution algorithms require a registration algorithm that can accurately determine relative sub-pixel motion between images. This thesis looks more specifically at a video sequence, where relative motion between image frames in the sequence is modelled as a simple planar translation. This section will therefore discuss the general super-resolution process as it applies to a video sequence.

2.2.1 Low-Resolution Frame Model. Digital cameras rely on an array of photo sensitive elements, called charged-coupled-devices, to capture an image. Each charge-coupled-device, or CCD, corresponds directly to a pixel value in an image. The resolution of the image depends upon the number and spacing of CCDs in the array. Because the number and spacing of the CCDs is finite, high frequency information is lost. Resolution is also affected by the lens and aperture of the camera, however this thesis is only concerned with restoring resolution lost by the CCD array.

Each low-resolution image (LR_i), created by the CCD array, is a decimated sampling of the high-resolution (HR) area it captured. At the pixel level the model becomes

$$LR_i(x, y) = De[HR(w, v)]. \quad (2.12)$$

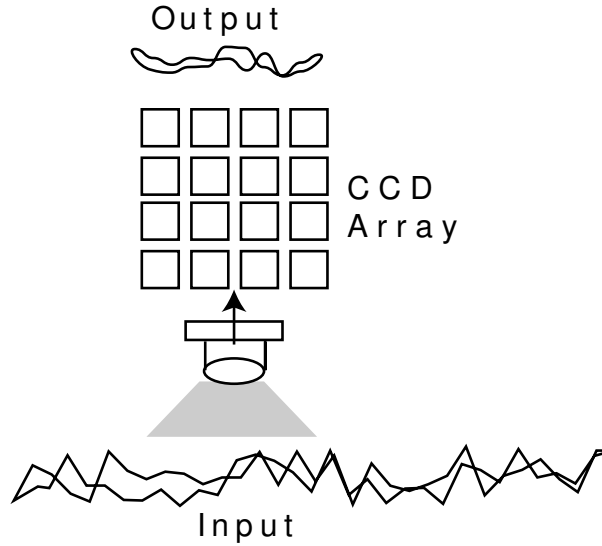


Figure 2.10: Band-Limited Finite CCD Array. Due to the finite number and spacing of CCD elements in the array, high frequency information is lost by the camera.

where $De(\cdot)$ is the decimation operator. The super-resolution algorithm determines $De(\cdot)$ and uses the LR_i to estimate HR .

Due to the motion of the camera, each image provides a different sampling of the same area. In order to restore the high-resolution information it is necessary to determine the relative motion between each low-resolution image. Once the motion is determined, the images are aligned and converted into a high-resolution frame. The next section explains this process as it applies to a video sequence.

2.2.2 Frame Alignment and Pixel Grid Conversion. The relative motion between frames in a video sequence depends upon the frame rate of the camera, and the velocity and direction of the camera's movements. It is assumed that all movements made by the camera are pure planar translations over any given area. Adjacent frames in a video sequence will overlap with much of the same information. A registration algorithm is required to determine the extent of this overlap. The overlapping portions of each frame are aligned with one another, which provides the multiple samples of the same area necessary for super-resolution.

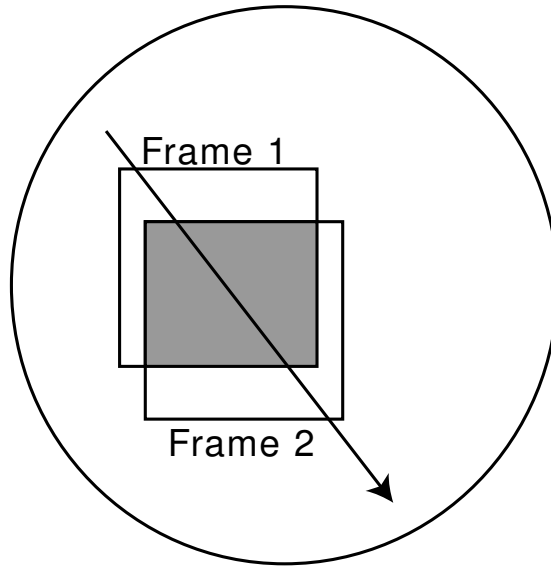


Figure 2.11: Overlapping Frames in a Video Sequence. Depending on the frame rate of the camera, and the velocity and direction of its movements, adjacent frames in a video sequence will overlap with much of the same information. The arrow indicates the direction of motion.

Depending on the algorithm, alignment may be determined at the high or low resolution pixel level. If frames are aligned at the low-resolution level, a pixel from one frame will align with a specific pixel in other overlapping frames. Each pixel is considered a sample of a set of high-resolution pixels. The number of high-resolution pixels in a set depends upon the degree of resolution improvement.

If frames are aligned at the high-resolution level, pixels may not perfectly align from one frame to another. This is referred to as a *sub-pixel shift*. The sub-pixel shift, which is utilized in Chapter 3, provides a more accurate alignment of the frames.

Once the registration algorithm has determined the alignment, each low-resolution frame is up-sampled into a high resolution estimate. The process of creating this estimate varies, depending upon the algorithm. The following section describes an effective method applied to video sequences. The method ties well into the algorithm developed in Chapter 3.

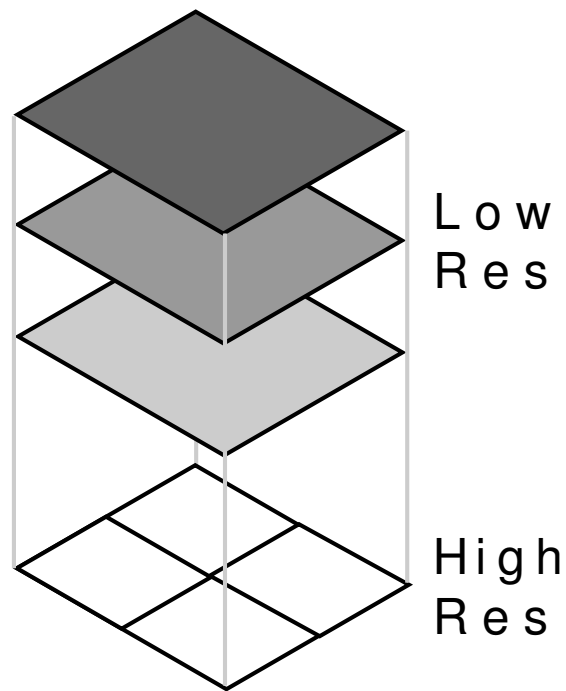


Figure 2.12: Low Resolution Pixel Alignment. Aligning frames at the low resolution pixel level makes each low-resolution pixel an estimate of a high-resolution set of pixels. This example assumes a 2x resolution improvement.

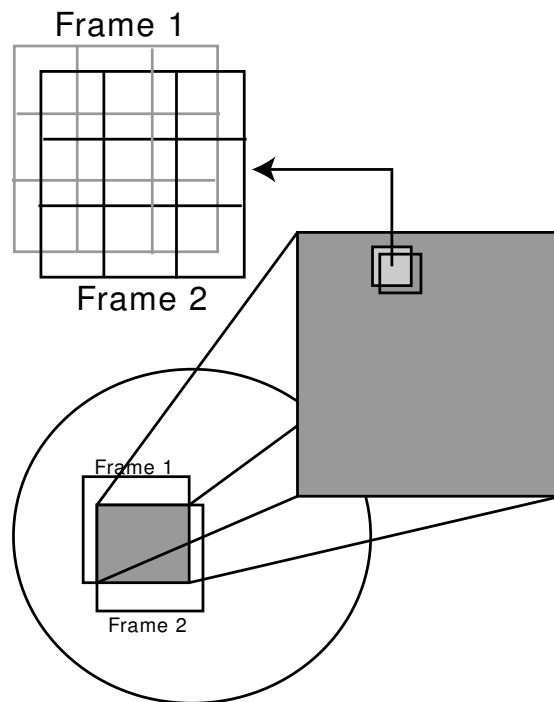


Figure 2.13: Sub-Pixel Shift. The two low-resolution frames are offset so that their pixels do not properly align. This is referred to as a sub-pixel shift.

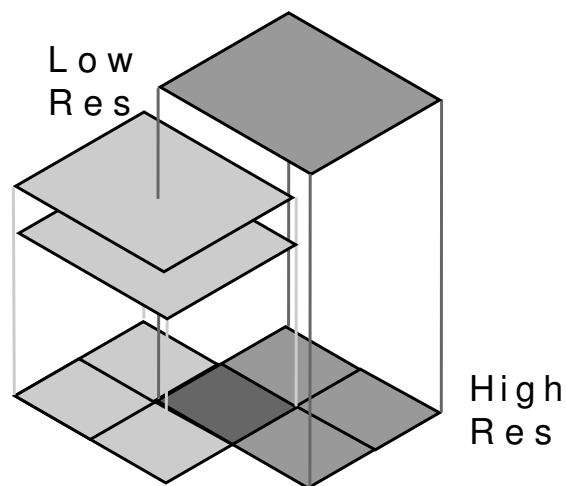


Figure 2.14: High-Resolution Pixel Alignment Due To Sub-Pixel Shift. When frames are aligned at the high-resolution level it is possible for low-resolution pixels to overlap and correspond to different high-resolution pixel sets.

2.2.3 Super-Resolution Estimation. A good super-resolution algorithm, which applies to video sequences, has been developed by Mendoza, Chen and Nakao [19]. It is based upon the super-resolution/optical flow algorithm proposed by Baker and Kanade [2]. The algorithm uses five frames in a video sequence to produce one high-resolution image.

The first step magnifies each frame to the desired high-resolution, using a tri-hybrid interpolation scheme [5, 18, 20]. The frames are then aligned with respect to the center frame. The first two frames are shifted forward, and the last two frames are shifted backward. A weighted average of the pixel values is then used to determine the final high-resolution image.

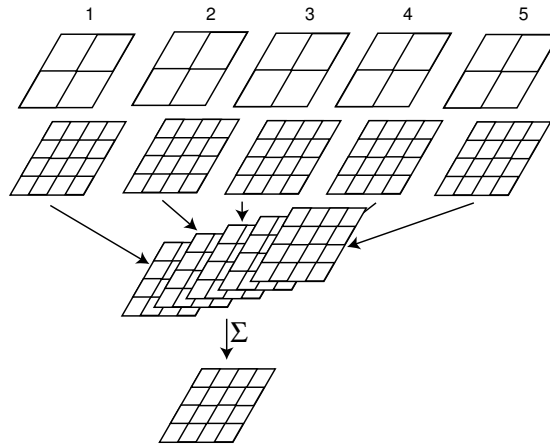


Figure 2.15: Mendoza Super-Resolution Algorithm. This super-resolution algorithm takes five frames from a video sequence and interpolates them to the desired resolution. The images are then aligned and a weighted sum is used to determine the pixel values of the final image.

This algorithm is similar to the one developed in Chapter 3. The images are aligned at the high-resolution level, and an interpolation scheme is used to determine the initial pixel values. However, the algorithm in Chapter 3 uses a wavelet based interpolation scheme and does not use a weighted sum to determine the final pixel values. The details are left to the next chapter.

2.2.4 Summary. This chapter discussed the wavelet transform and super-resolution fundamentals. The wavelet transform was shown to provide a representation of an image that is useful for applications such as image registration and interpolation. Super-resolution was shown to restore high-frequency information lost by the CCD array of the camera.

Image registration is a key first step to super-resolution. Many good registration algorithms use wavelet-based techniques. This thesis explores the possibility of extending the use of the wavelet transform to super-resolution. Using wavelets for both super-resolution and image-registration may prove to be more computationally efficient, and produce better quality images. Chapter 3 explores this idea in detail.

III. Methodology

The super-resolution algorithm developed in this chapter uses a wavelet-based interpolation scheme to create one high-resolution image from two adjacent low-resolution frames. It is assumed that a registration algorithm exists which can take a video sequence and determine the pixel shift between adjacent frames at the sub-pixel level. The algorithm also assumes a specific sub-pixel shift between the frames. All test data applied to the super-resolution algorithm has this shift. Therefore, before the algorithm can be discussed it is first necessary to understand the specific sub-pixel shift of the test data.

3.1 The Test Data - Half Pixel Shifts

The test data created for the super-resolution algorithm is comprised of two frames that are at a half-pixel shift in the x and y direction relative to one another. This shift was chosen because it creates a commonly used sampling pattern called quincunx sampling. Quincunx sampling is best described as only taking pixels in an image with purely even, and purely odd, valued indices, as shown in Figure 3.1. It is an appropriate starting point because if the algorithm is shown to be effective for this shift, it may be extended to other arbitrary sub-pixel shifts.

The shift simulates the diagonal motion of a camera over an area (See Figure 2.11). To create the data a 256x256 pixel image is split into two 128x128 pixel images. This creates the quincunx sampling described earlier. It is assumed that as the camera moves diagonally over an area, the CCD array of the camera has a spacing and size that samples the area in this half-pixel manner. Unlike other super-resolution algorithms, where one low-resolution pixel corresponds to a set of high-resolution pixels, each low-resolution pixel is assumed to correspond to only one high-resolution pixel. Figure 3.2 illustrates this difference. This is based on the fact that we are only dealing with resolution lost by the finite number and spacing of the CCD array. Therefore, no low pass filter was applied to the data. The pixels

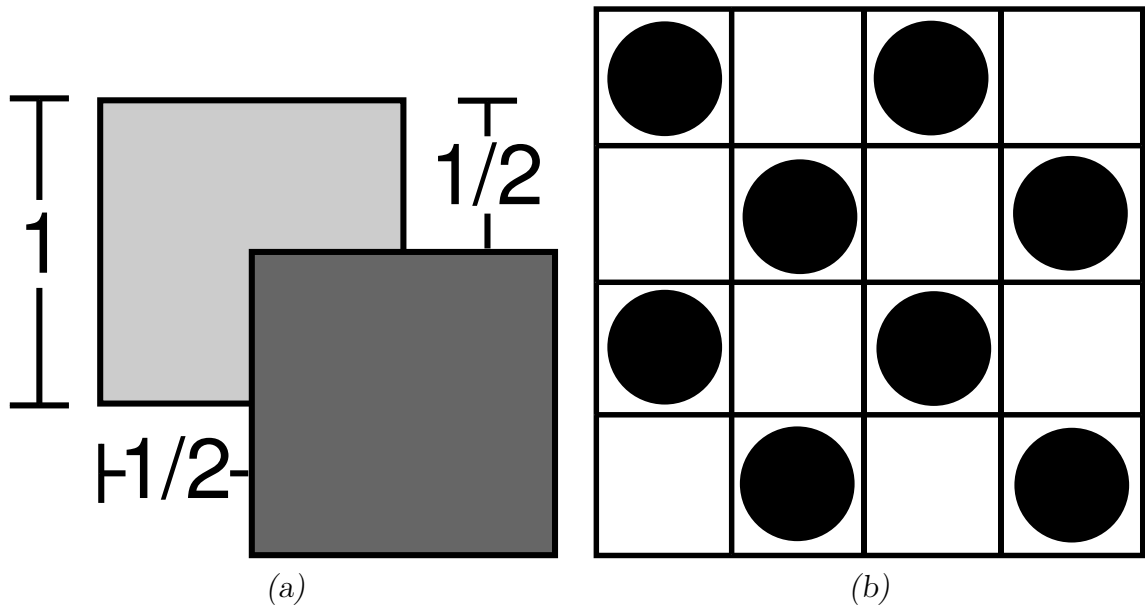


Figure 3.1: Half Pixel Shift and Quincunx Sampling. (a) A half-pixel shift in the x and y direction between 2 pixels from adjacent frames. (b) Quincunx Sampling.

from the low-resolution frames are assumed to be actual pixels in the high-resolution frame. The idea is to recreate the 256x256 pixel image from the two 128x128 frames. Figures 3.3 shows how a 6x6 pixel image is split into two 3x3 pixel images at the desired half-pixel shift.

3.2 The Wavelet Super-Resolution Algorithm

The wavelet super-resolution algorithm deals specifically with two frames separated by a half-pixel shift. The two frames are combined into a quincunx sampling grid and rotated to optimize pixel correlation. The redundant discrete wavelet transform is then applied to the rotated pixel set, and the high-resolution pixels are interpolated. The inverse redundant discrete wavelet transform is then applied, and the image is rotated back to its original orientation. A few post processing steps are then applied to the image to create the final high-resolution image.

3.2.1 Pre-Wavelet Domain Operations. Before the wavelet transform can be applied to the two low-resolution frames they must be combined and adjusted to

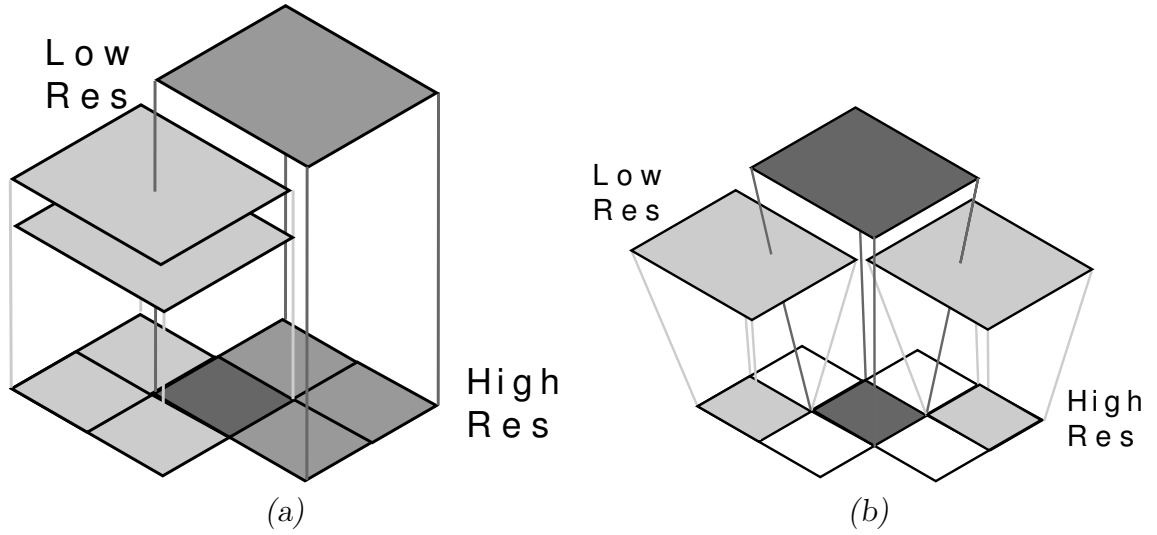


Figure 3.2: *Traditional vs. My Super-Resolution Pixel Association.* (a) A low-resolution pixel is normally assumed to correspond to a set of high-resolution pixels. (b) Since this thesis deals only with resolution loss by the CCD array of the camera, a low-resolution pixel only corresponds to one high resolution pixel.

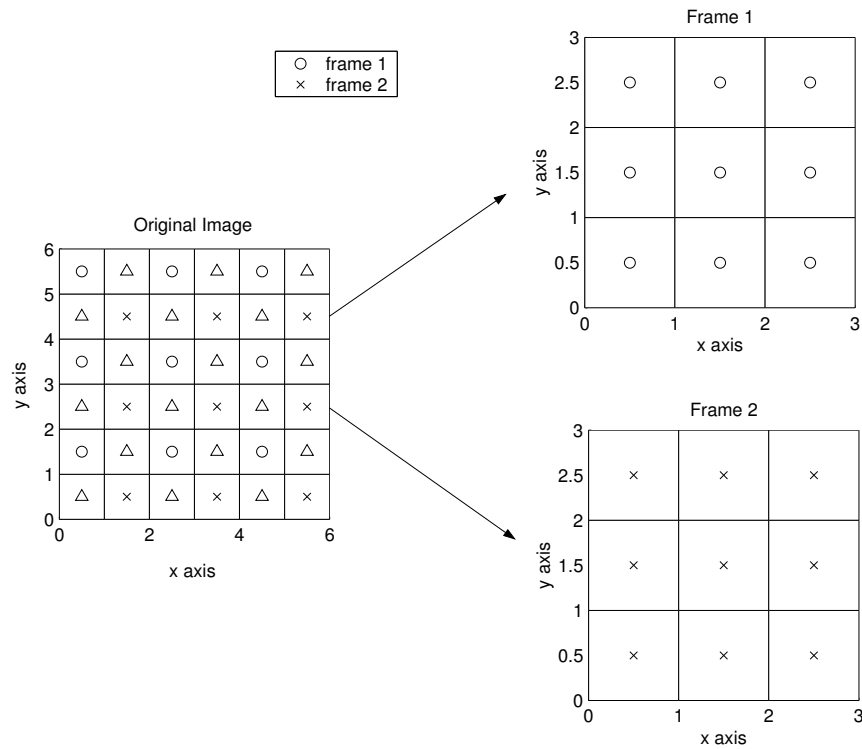


Figure 3.3: *Half Pixel Shift Application.* A 6x6 pixel image is split into two 3x3 pixel images which are at a half pixel shift in the x and y direction from each other. They have a resolution which is half that of the original 6x6 image.

produce the highest quality result. These operations consist of combining, rotating, and up-sampling the frames.

3.2.1.1 Combining Frames. The first step in the super-resolution process is to combine the two test frames. As explained in Section 3.1, each low-resolution pixel corresponds to one high-resolution pixel. Placing the low-resolution pixels into the high-resolution grid becomes a matter of knowing how the two frames are positioned relative to one another. Since the shift is known to be a half pixel in the x and y direction, the process is straightforward. The high-resolution grid will be designated as $H(x_H, y_H)$, and the frames as $F_1(x_1, y_1)$ and $F_2(x_2, y_2)$. The process that transforms $F_1(x_1, y_1)$ and $F_2(x_2, y_2)$ into $H(x_H, y_H)$ will be designated as $Tr(\cdot)$. This process is shown in Figure 3.5(a) where a 1 represents Frame 1 pixels and a 2 represents Frame 2 pixels.

$$H(x_H, y_H) = Tr[F_1(x_1, y_1), F_2(x_2, y_2)] \quad (3.1)$$

$$x_1 = 2x_H - 1 \quad y_1 = 2y_H - 1 \quad (3.2)$$

$$x_2 = 2x_H \quad y_2 = 2y_H \quad (3.3)$$

The result is a quincunx sampling of the desired high-resolution image. Figure 3.4 illustrates this, as it is applied in the algorithm. The complete high-resolution image is created by interpolating the missing pixels in the wavelet domain, using the 2D redundant discrete wavelet transform. However, before the quincunx sampled image is converted to the wavelet domain, a rotation must be applied.

3.2.1.2 Quincunx Image Rotation. Chapter 2 explained that the 2D wavelet transform works strictly along the columns and rows of an image. Because of this, applying the wavelet transform to the current quincunx sampled image would be the same as applying the wavelet transform to each frame individually. No in-

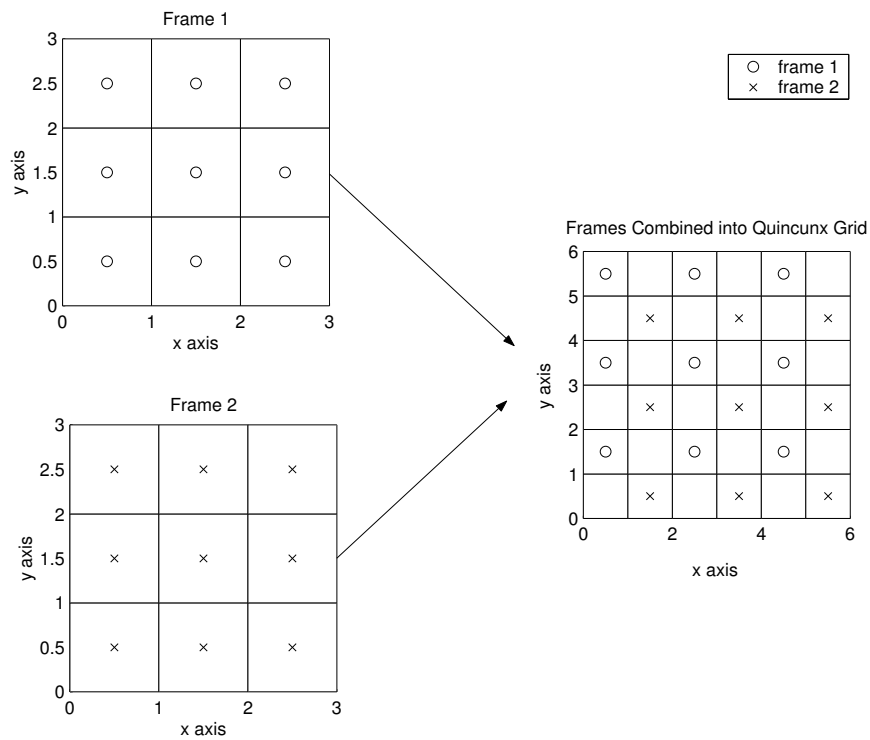


Figure 3.4: Combining Test Frames Into High Resolution Grid. This example combines the test frames shown in Figure 3.3 into the original high resolution grid.

formation about the relationship between the two frames would be introduced. A 45-degree rotation is needed to exploit this relationship. Visually, a 45-degree rotation is easy to realize (See Figure 3.5(b)), however, this changes the index values of each pixel. The indices of each pixel have to be transformed from the quincunx high-resolution grid to the rotated high-resolution grid (See Figure 3.5(c)). Let Rot denote the 45-degree rotation transform, and H_{rot} denote the rotated grid, then

$$H_{rot}(x_{rot}, y_{rot}) = Rot[H(x_H, y_H)] \quad (3.4)$$

or

$$H_{rot}(x_{rot}, y_{rot}) = H(Rot[x_H], Rot[y_H]). \quad (3.5)$$

Before the wavelet transform can be applied to the rotated image, it must be up-sampled. The following section explains why.

3.2.1.3 Up-sampling. The rotated image must be up-sampled by a factor of 2 to create the space required for each missing pixel. The wavelet transform filters must also be up-sampled by 2. Up-sampling the image and the filters ensures that each row and column of the final high-resolution image was operated on separately by the wavelet transform filters. This is especially important when applying the IRDWT, to ensure that the pixel values from the original two frames are not changed. Figure 3.5(d) illustrates this up-sampling, where an M represents the location of missing high-resolution pixel:

$$H_R(x_R, y_R) = Up[H_{rot}(x_{rot}, y_{rot})], d \quad (3.6)$$

$$x_R = 2x_{rot} \quad y_R = 2y_{rot} \quad (3.7)$$

Figure 3.5 shows the combining, rotating, and up-sampling operations performed on the two low-resolution frames prior to being transformed into the wavelet domain.

The frames are now ready to be transformed into the wavelet domain. In the wavelet domain missing pixel values are located and interpolated. The next section discusses these operations.

3.2.2 Wavelet Domain Operations. The uniqueness of this super-resolution algorithm lies in its application of the wavelet transform. It utilizes the parsimonious clustering property of the wavelet transform to interpolate missing high resolution pixels. Chapter II described how each subband in the wavelet domain emphasizes different edges. By interpolating missing pixels in the wavelet domain, it is theorized that edges will be more accurately reconstructed and preserved.

In order to interpolate the missing pixels their locations must be determined. In the wavelet domain, missing pixels correspond to missing coefficients. By determining the location of, and interpolating the missing coefficients, the complete high-resolution image can be constructed. The process of determining the missing coefficient locations is not pertinent to this discussion and is explained in Appendix A.

3.2.2.1 Interpolation. The missing pixel coefficient is calculated using the four coefficients closest to the missing pixel location. Based on the clustering property, the coefficients nearest the missing coefficient will best approximate its value. Missing coefficients are designated with an M and known coefficients are denoted C . C_i for $i = 1, 2, 3, 4$ denote the four closest coefficients, as shown in Figure 3.6. These four coefficients may correspond to a number of different image phenomena, as shown in Figure 3.7. They may represent a line, an edge, or an area with neither a line nor an edge. Furthermore, the line or edge may be horizontal, diagonal, or vertical. If the missing coefficient exists in an area with no line or edge,

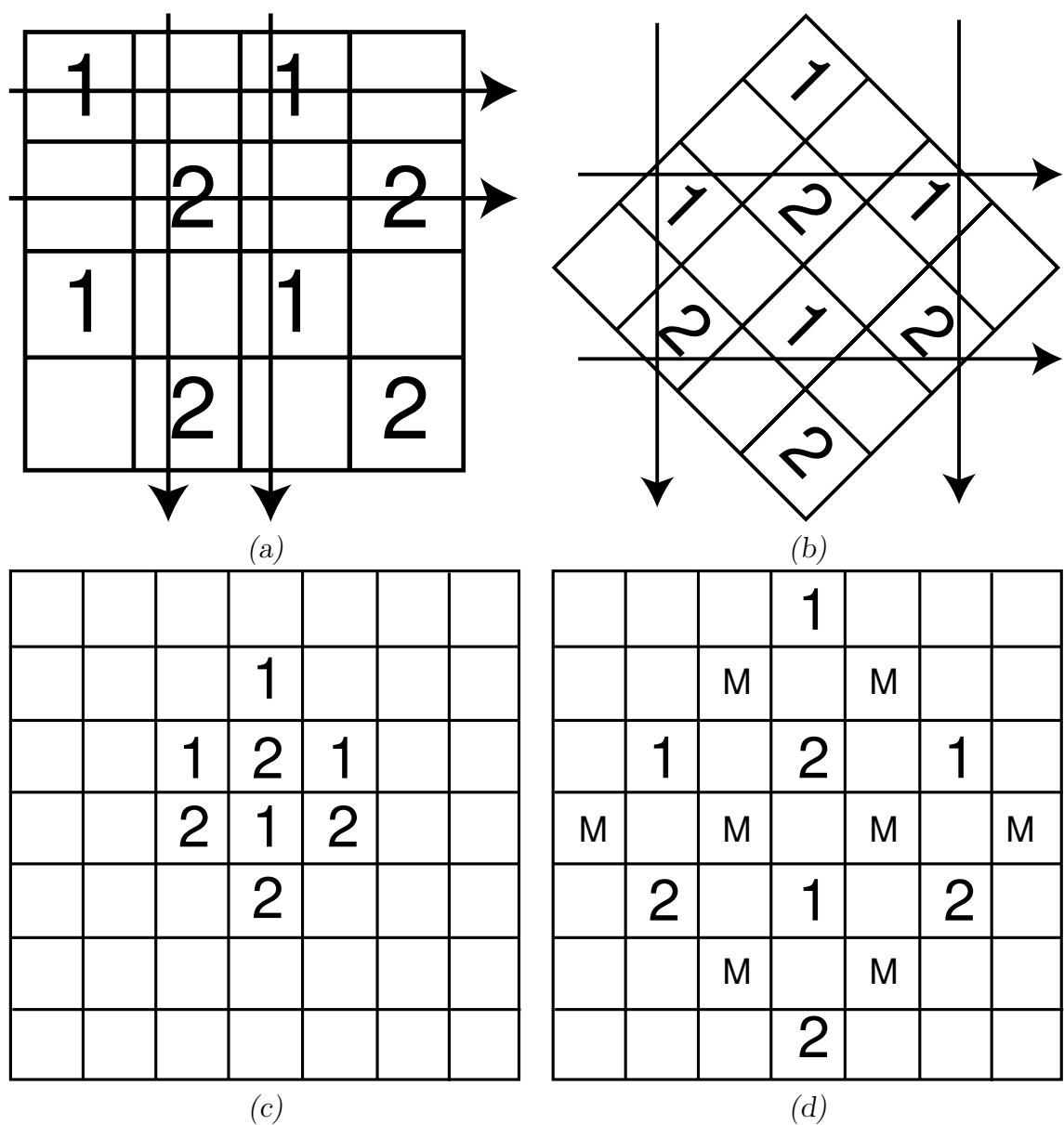


Figure 3.5: Pre-Wavelet Domain Operations. (a) non-optimal “quincunx” orientation (b) 45-degree rotation which optimizes use of the wavelet transform. Arrows indicate the movement of the wavelet transform filters across the image. (c) The indices of the pixels must be transformed into a new rotated grid. (d) The rotated pixels are up-sampled by a factor of two to make room for the missing pixels and to ensure that each row and column of the final high-resolution image is operated on separately by the wavelet transform filters.

C		C		C		C
	M		M		M	
C		C ₁		C ₂		C
	M		M		M	
C		C ₄		C ₃		C
	M		M		M	
C		C		C		C

Figure 3.6: Missing Coefficient and the Four Nearest Neighbors. A C represents the value of the known coefficients and M represents an unknown coefficient. C_i for $i = 1, 2, 3, 4$ are the four nearest known coefficients, used to determine the value of M.

it may be interpolated as an average of the four surrounding coefficients:

$$M = \frac{C_1 + C_2 + C_3 + C_4}{4}. \quad (3.8)$$

However, in the presence of an edge or a line the coefficient must be calculated using the coefficients associated with the edge or line to which it belongs. When the missing coefficient corresponds to its own line, the value is impossible to accurately determine. This begs the question of how best to determine the group of coefficients the missing coefficient belongs to.

If the missing coefficient is associated with a group of two other coefficients the resulting value will be arbitrarily skewed towards one edge or line. If it is associated with the entire group of four, the value will not accurately represent any edge or line. Therefore, the best 3 of 4 pixels is chosen to determine the value. It is assumed that the more closely related the surrounding pixels are, the more likely the missing pixel is to associate with them. The standard deviation of every possible group of

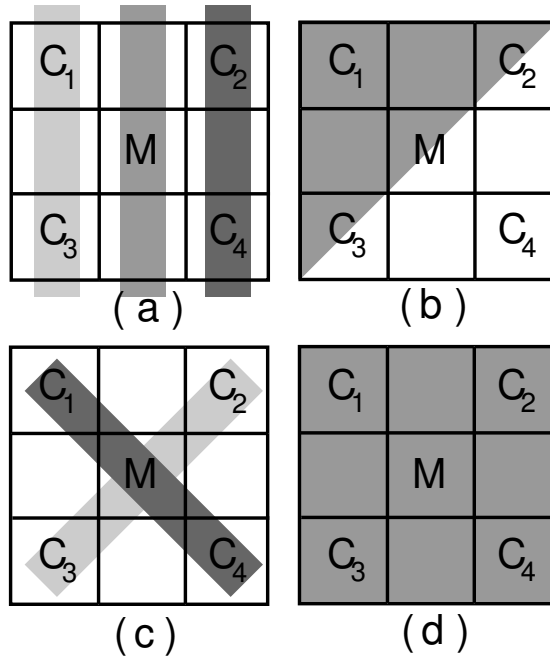


Figure 3.7: Pixel Groupings. (a) A line or an edge occurring above or below the missing pixel. (b) An edge occurring towards one side of the missing pixel. (c) A line occurring on a diagonal with the missing pixel. (d) The pixel is in a flat area. Note that it is impossible to determine a line that is created solely by missing pixels.

three is calculated to determine this *best* group. The average of the group with the lowest standard deviation is used as the value of the missing pixel. This is shown in Equations 3.9 through 3.14.

$$G_1 = \{C_1, C_2, C_3\} \quad (3.9)$$

$$G_2 = \{C_2, C_3, C_4\} \quad (3.10)$$

$$G_3 = \{C_3, C_4, C_1\} \quad (3.11)$$

$$G_4 = \{C_4, C_1, C_2\}, \quad (3.12)$$

$$k = \arg \min_i [std(G_i)], \quad (3.13)$$

$$M = mean[G_k]. \quad (3.14)$$

Once this process is completed for each missing pixel location, the inverse redundant wavelet transform (IRDWT) is applied to the rotated image, with the necessary up-sampling of the synthesis filters.

Now the new rotated high-resolution image is rotated back into its original orientation. Recall that Equation 3.5 represented the operation that transformed the quincunx high-resolution grid into the rotated high-resolution grid. The inverse of this operation is now performed on the new image:

$$H(x_H, y_H) = Rot^{-1}[H_{rot}(x_{rot}, y_{rot})]. \quad (3.15)$$

Since the missing pixels have been interpolated, the holes of the quincunx sampling grid are filled and the complete high-resolution image is produced. This process is illustrated in Figure 3.8.

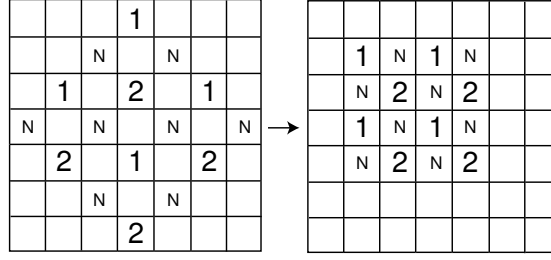


Figure 3.8: Re-Rotation of High-Resolution Image. After the missing pixels have been interpolated (Labeled N) and the IRDWT is applied to the complete set of image coefficients, the image is re-rotated into its original orientation.

Though the high-resolution image has all its pixel values, two more steps must be completed to ensure that the new pixels are as accurate as possible. The wavelet transform is used once more to *clean-up* the pixel values.

3.2.3 Post Wavelet Domain Processing. The first post processing step adjusts the range of the pixel values. Due to interpolation error, when the IRDWT is applied to the interpolated pixels, some values may be below 0 or above 255. Since this is outside the possible range of values, all pixels outside of this range are truncated to 0 or 255 accordingly.

The non-rotated RDWT of this new super-resolution image exposes interpolation errors. As a result of the interpolation scheme used to create the new coefficients, lines and edges associated solely with these new coefficients are not emphasized. Since new pixel values are oriented diagonally with each other, this means that diagonal lines and edges are poorly reconstructed and horizontal and vertical lines are over-emphasized. Figure 3.9 illustrates this problem as it appears in the HH subband.

To compensate for this problem the image is brought back into the wavelet domain, without rotation or up-sampling. Chapter 2 explained that the HH subband emphasizes diagonal edges in an image. However, the HH subband of the new high-resolution image shows an emphasis of some vertical and horizontal edges too, as shown in Figure 3.10. A two-step filtering process is applied to eliminate the

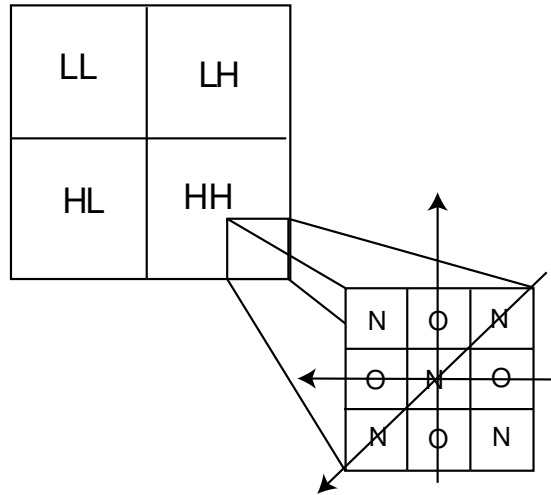


Figure 3.9: Pixel Correlations. Due to the interpolation process, new pixel values (N) poorly reconstruct the edges or lines only associated with them. Vertical and horizontal edges and lines, which are a combination of new and old (O) pixels, are still accurately reconstructed.

erroneous edges. The first step applies the analysis low-pass filter to the rows and columns of the HH subband. This subdues all edges in the HH subband. Since only diagonal edges are to be emphasized in this subband, the second step applies the high-pass synthesis filter to the rows and columns of the new subdued HH subband. This enhances the diagonal edges, and keeps the horizontal and vertical edges subdued. Figure 3.11 shows this process, and the results are illustrated in Chapter 4. The IRDWT is then applied to the adjusted wavelet transform to create a new corrected high-resolution image.

The filtering process eliminates erroneous edges, and enhances diagonal ones, but it also alters the values of the original pixels from the two low-resolution frames. The only pixels that needed to be altered were the new interpolated pixels. Therefore, only the altered new pixels are used in the final high-resolution image. The original pixels are left unaltered.

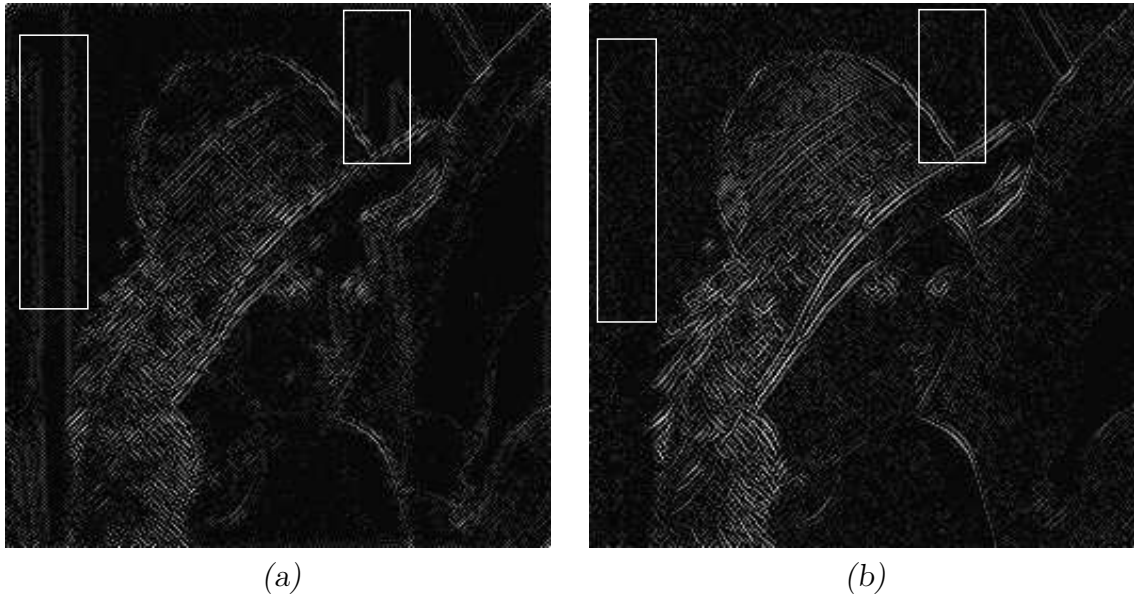


Figure 3.10: Erroneous Edges in HH Subband. (a) This is the HH subband of high-resolution image created by the super-resolution algorithm. Notice the erroneous vertical lines highlighted by the boxes (b) This is the HH subband of the original image for comparison to show the absence of the vertical lines.

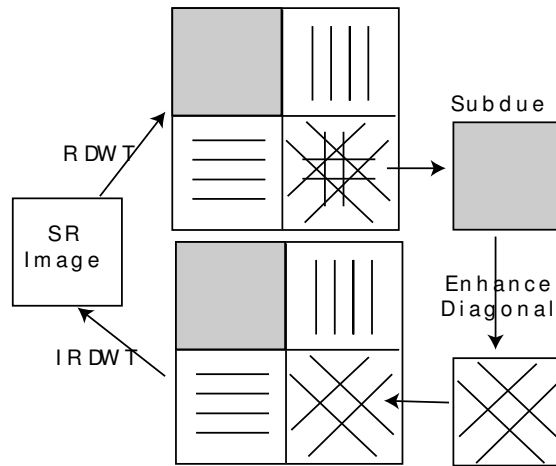


Figure 3.11: Correcting the HH Subband. Erroneous vertical and horizontal lines are removed from the HH Subband

3.3 *Summary*

The wavelet super-resolution algorithm developed here creates one high-resolution image from two low-resolution frames, oriented at a specific half pixel shift. The frames are placed into a high-resolution frame according to this shift, then rotated and up-sampled to most effectively utilize the wavelet transform. The RDWT is applied and the missing pixel coefficients are interpolated. The interpolation process adaptively chooses 3 or 4 of the coefficients to estimate the value of the missing coefficients. The IRDWT is then applied to the final set of coefficients to create the new high-resolution image. The HH subband of this high-resolution image is adjusted to eliminate erroneous edges, and enhance diagonal edges. As will be shown in Chapter 4, this unique wavelet-based process effectively reconstructs the original high-resolution image from the low-resolution frames.

IV. Results

An analysis of algorithm performance is presented here. The resolution reconstruction performance of the algorithm is tested in a variety of different image environments, using six different image. Various wavelets are tested in the algorithm to determine which wavelets perform best. Comparisons are also made between the applied wavelet technique, and traditional techniques. The mean square error and L-infinity error for each technique is then discussed to provide a quantitative performance analysis. A portion of the chapter is devoted to analyzing the improvement made by the post processing techniques described in Chapter 3.

4.1 The Test Images

Six distinctly different images are chosen to test the performance of the super-resolution algorithm. An image with different sized circles is used to test the reconstruction performance of rounded edges (Figure 4.1(a)), while an image with different sized rectangles is used to test the reconstruction of vertical and horizontal edges (Figure 4.1(c)). Finally, an image called *pinwheel* will test reconstruction of diagonal edges (Figure 4.2(a)). These three are purely test images, used to separately analyze the reconstruction of shapes and edges of different sizes. Three other real world images are chosen to test the real world application of the algorithm.

Each real world image was chosen for its differing degree of detail. The classic Lenna portrait is an image near the object of interest with many fine details (Figure 4.2(c)), while the Cameraman image is a little farther away with both fine detail and coarse background information (Figure 4.3(a)). Finally, an aerial view image will be used to test the reconstruction of purely coarse information (Figure 4.3(c)).

4.2 The Wavelets

Three wavelet systems are chosen for analysis. The first is the classic Haar wavelet system, which contains length 2 orthogonal filters. This is the simplest

Haar

A_L :	0.7071	0.7071
A_H :	-0.7071	0.7071
S_L :	0.7071	0.7071
S_H :	0.7071	-0.7071

Daubechies 9,7

A_L :	0.0378	-0.0238	-0.1106	0.3774	0.8527	0.3774	-0.1106	-0.0238	0.0378
A_H :		-0.0645	0.0407	0.4181	-0.7885	0.4181	0.0407	-0.0645	
S_L :		-0.0645	-0.0407	0.4181	0.7885	0.4181	-0.0407	-0.0645	
S_H :	-0.0378	-0.0238	0.1106	0.3774	-0.8527	0.3774	0.1106	-0.0238	-0.0378

Daubechies 7,9

A_L :		-0.0645	-0.0407	0.4181	0.7885	0.4181	-0.0407	-0.0645	
A_H :	-0.0378	-0.0238	0.1106	0.3774	-0.8527	0.3774	0.1106	-0.0238	-0.0378
S_L :	0.0378	-0.0238	-0.1106	0.3774	0.8527	0.3774	-0.1106	-0.0238	0.0378
S_H :		-0.0645	0.0407	0.4181	-0.7885	0.4181	0.0407	-0.0645	

Table 4.1: Wavelets Used for Testing.

wavelet system and provides a good benchmark from which to compare the performance of other systems. The Daubechies 9, 7 and 7, 9 wavelet systems are the final two used for analysis. They are chosen due to their excellent localization properties. They are biorthogonal odd length filters, which provide a more parsimonious representation of the signal. From the discussion in Chapter 2, this should improve interpolation performance. Moreover, Brown [3] and Manfra [17] have already shown the Daubechies 9, 7 wavelet system to have excellent image registration performance. This performance is extended by this thesis into super-resolution. The difference between the 9, 7 and 7, 9 wavelets has to do with the length of the high- and low- pass analysis filters, as shown in Table 4.1. This affects the smoothness of the analysis and synthesis filters. Typically, smoother synthesis filters provide better reconstruction, however it will be shown that this is not the case for the super-resolution algorithm.

4.3 The Comparisons

Three different comparison experiments are used in this chapter. The test data is the same for all experiments. Each original 256×256 pixel image is split into

two 128×128 pixel frames, at the specific half-pixel shift described in Section 3.1. These frames are applied to the super-resolution algorithm as well as to three other interpolation techniques. The first experiment compares the performance of the different wavelets, to determine which wavelet performs best. The next experiment tests the best performing wavelet against three other interpolation techniques. The final experiment compares the error of the final high-resolution image before and after the post-processing techniques described in Chapter 3. Each experiment uses mean square error (MSE), and L-infinity error, to quantitatively analyze performance. MSE is used because it is the typical error power measure used as industry standard. The L-infinity error is used because it is based on maximum deviation, and is more closely related to how the human visual system works.

4.4 Analysis

The analysis of the three experiments is broken into three parts. The first part provides a qualitative analysis of the images. Each image is compared as a whole, and then a portion of each image is selected to more clearly illustrate reconstruction performance. Following this is a quantitative analysis of the MSE and L-infinity errors.

4.4.0.1 Qualitative analysis. All of the super-resolution images shown in this section were produced using the Daubechies 7,9 wavelet, as it was the best performing wavelet. It should be noted that the other wavelets did produce comparable results. Figures 4.1 through 4.3 compare each image as a whole, while Figures 4.4 through 4.13 analyze a specific portion of each image.

In the circles and rectangles images of Figure 4.1(a) and (b) an offset value is obvious. This is because the original images are binary, and the super-resolution algorithm did not reconstruct each missing pixel at the 0 and 255 binary values. Aliasing is noticeable in the smaller rectangles of Figure 4.1(a). All rectangles are distinguishable from one another, with the exception of the four smallest rectangles

in the upper left-hand corner of the image. The algorithm did an excellent job of reconstructing the pinwheel image, as illustrated in the comparison of Figures 4.2(a) and (b). The only noticeable difference is some aliasing at the center of the image, where the highest spatial frequencies occur.

Each real world image was accurately reconstructed, as shown in Figure 4.2(a) through 4.3. Differences in the Lenna image are noticeable along the rim of the hat, and in the details of the feather. In the Cameraman image the only noticeable loss in detail is in the fine lines of the camera. Notice how accurately the tripod legs are reconstructed. The aerial image has no noticeable loss in detail, but does illustrate a slight offset from the original image. The reason for this offset is similar to the offset reason for the binary images. The algorithm did not exactly determine the pixel values for the coarse areas.

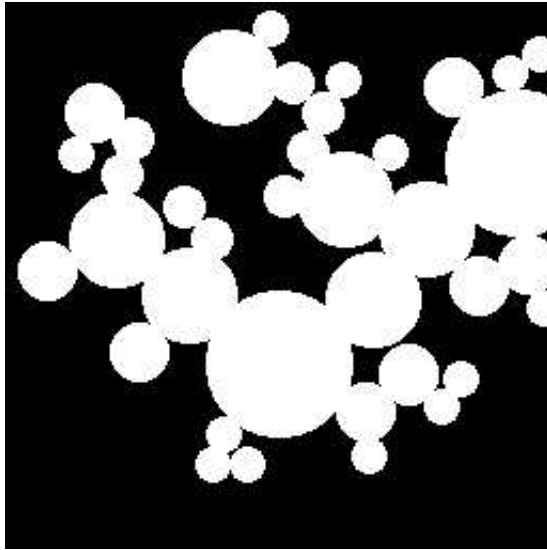
Figures 4.4 through 4.9 analyze a specific portion of each image. Each low-resolution frame is illustrated in these figures, along with the original and the super-resolution image. The offset is again apparent in the Circles and Rectangles image, but the edges of each image are accurately reconstructed. The individual frames of the Rectangles image show irregular spacing and size of the smaller rectangles in the image, but the super-resolution algorithm is able to fairly accurately restore the spacing and size. Each rectangle is individually distinguishable, and the spacing is correct. In the Lenna image, shown in Figure 4.7, a considerable amount of detail is lost in the feather and hat of the individual frames. The super-resolution algorithm does an excellent job of restoring this detail. The Cameraman image does perhaps the best job of illustrating the algorithm's excellent image reconstruction performance (Figure 4.8). The details of the face, camera, tripod and background are greatly distorted in each frame, and still properly restored by the super-resolution algorithm. In the aerial view of Figure 4.9 individual cars in the parking lot of each frame are difficult to distinguish. In the super-resolution image, however, the cars are immediately distinguishable.

Figures 4.10 through 4.13 compare the super-resolution reconstructed image to reconstructed images using cubic spline interpolation. Cubic spline is used for comparison because it provides the best error results, as shown in Tables 4.3 and 4.5. In each image it is obvious that the wavelet based super-resolution image has better reconstruction performance. This is most apparent in the reconstruction of the rectangles, shown in Figure 4.11. The smaller rectangles are poorly reconstructed in the cubic interpolated image. In the Lenna and Cameraman images, Figures 4.12 and 4.13, the reconstruction of the edges in each image is much more accurate in the wavelet super-resolution image.

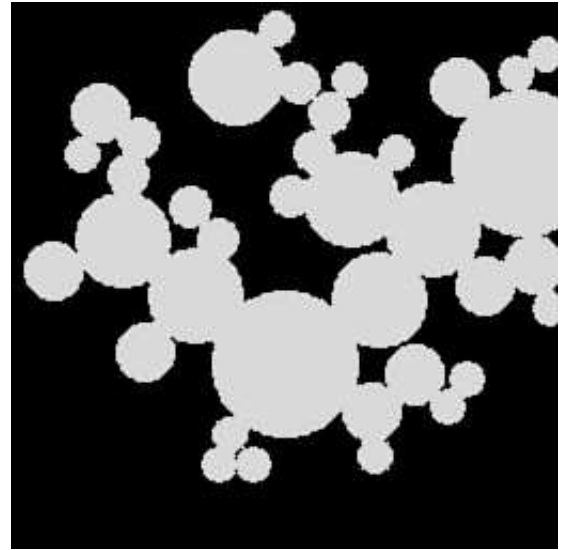
Figure 4.14 shows the before and after images of the post-processing portion of the algorithm. Though substantial differences are not noticeable, the dark diagonal lines of the after image do have a smoother appearance. The benefit of this post-processing is best seen in the error results in Tables 4.6 and 4.7.

4.4.0.2 Quantitative Analysis. Tables 4.2 through 4.7 show the mean square error and L-infinity error of each experiment. Tables 4.2 and 4.3 show that the Daubechies 7,9 wavelet system performs the best, in terms of mean square error. The difference between the Daubechies 7,9 and the traditional interpolation schemes is substantial. In terms of L-infinity error, all of the tested techniques performed roughly the same on the binary images. In Table 4.4, the Daubechies 7,9 performed better on the pinwheel and Lenna image, while the Daubechies 9,7 performed better on the aerial and cameraman images. This suggests that the Daubechies 7,9 will perform better on images with fine detail, and the Daubechies 9,7 wavelet will perform better on images with coarse information. Overall, the wavelet-based super-resolution algorithm performed significantly better than the other test interpolation schemes.

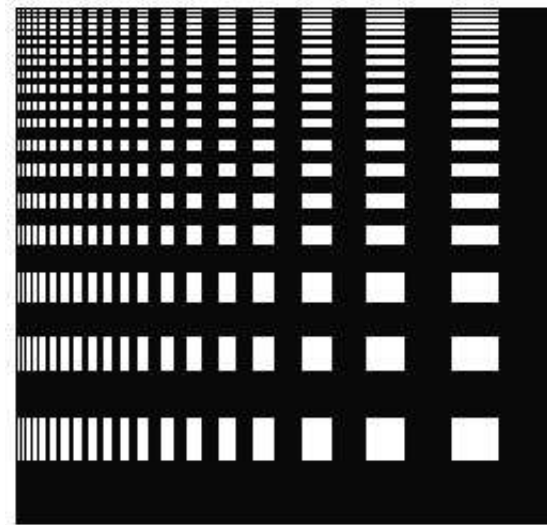
The before and after errors of Tables 4.6 and 4.7 were generated using the Daubechies 7,9 wavelet system. The error improvement of the post-processing is



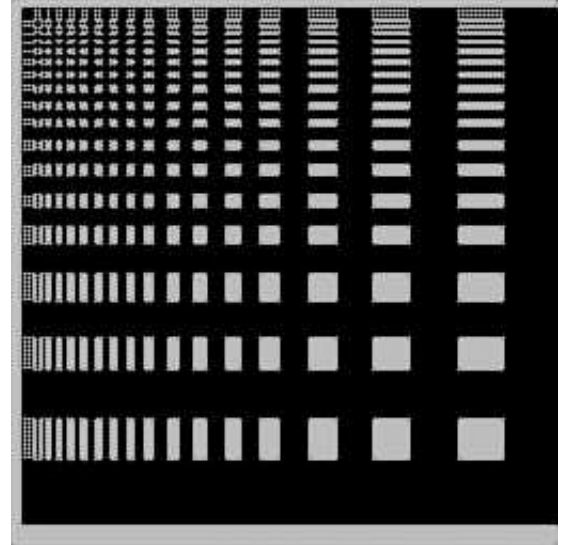
(a)



(b)

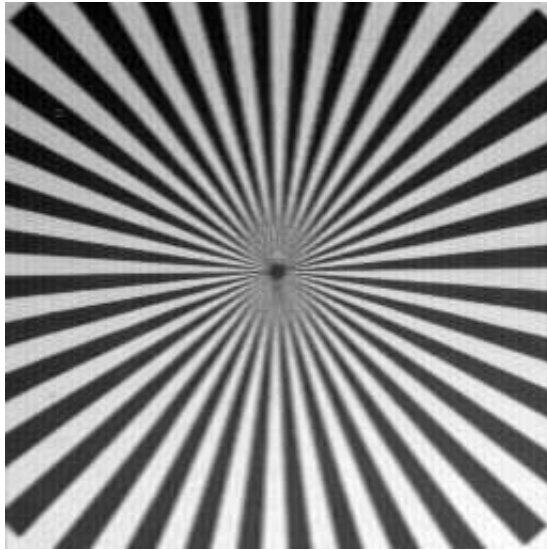


(c)

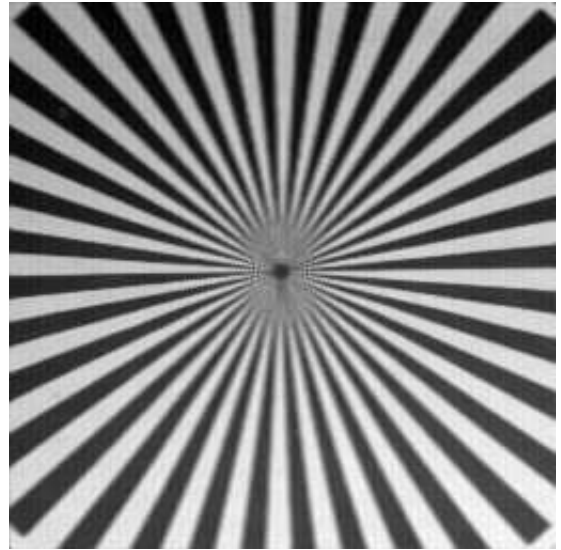


(d)

Figure 4.1: Original and Super-Resolution Test Images. (a) Original Circles(binary) (b) Super-Resolution Circles (c) Original Rectangles(Binary) (d) Super-Resolution Rectangles



(a)



(b)



(c)



(d)

Figure 4.2: Original and Super-Resolution Test Images. (a) Original Pinwheel (b) Super-Resolution Pinwheel (c) Original Lenna (d) Super-Resolution Lenna



(a)



(b)

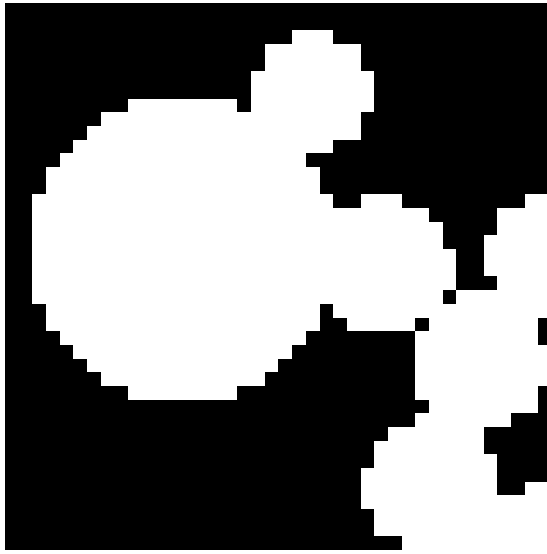


(c)



(d)

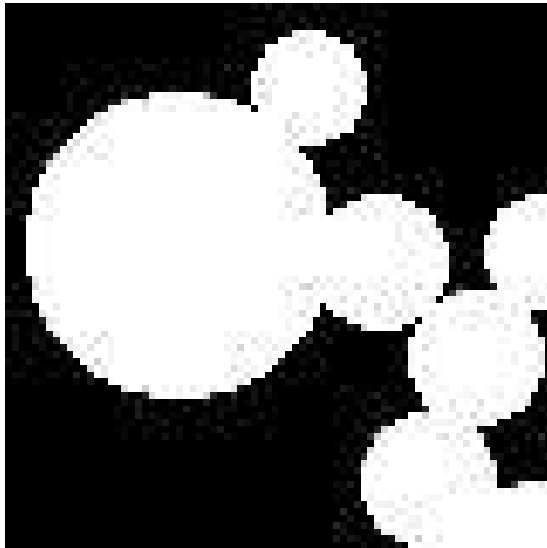
Figure 4.3: Original and Super-Resolution Test Images. (a) Original Cameraman (b) Super-Resolution Cameraman (c) Original Aerial (d) Super-Resolution Aerial



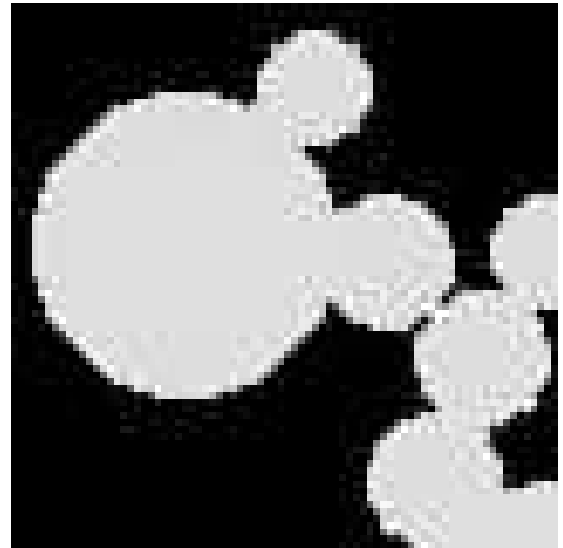
(a)



(b)



(c)



(d)

Figure 4.4: Circles Comparison. (a) 1st Frame. (b) 2nd Frame. (c) Original Figure. (d) Super-Resolution Image. The edge of each circle is slightly blurred, but still properly reconstructed. The gray offset is due to interpolation error, but no detail is lost.

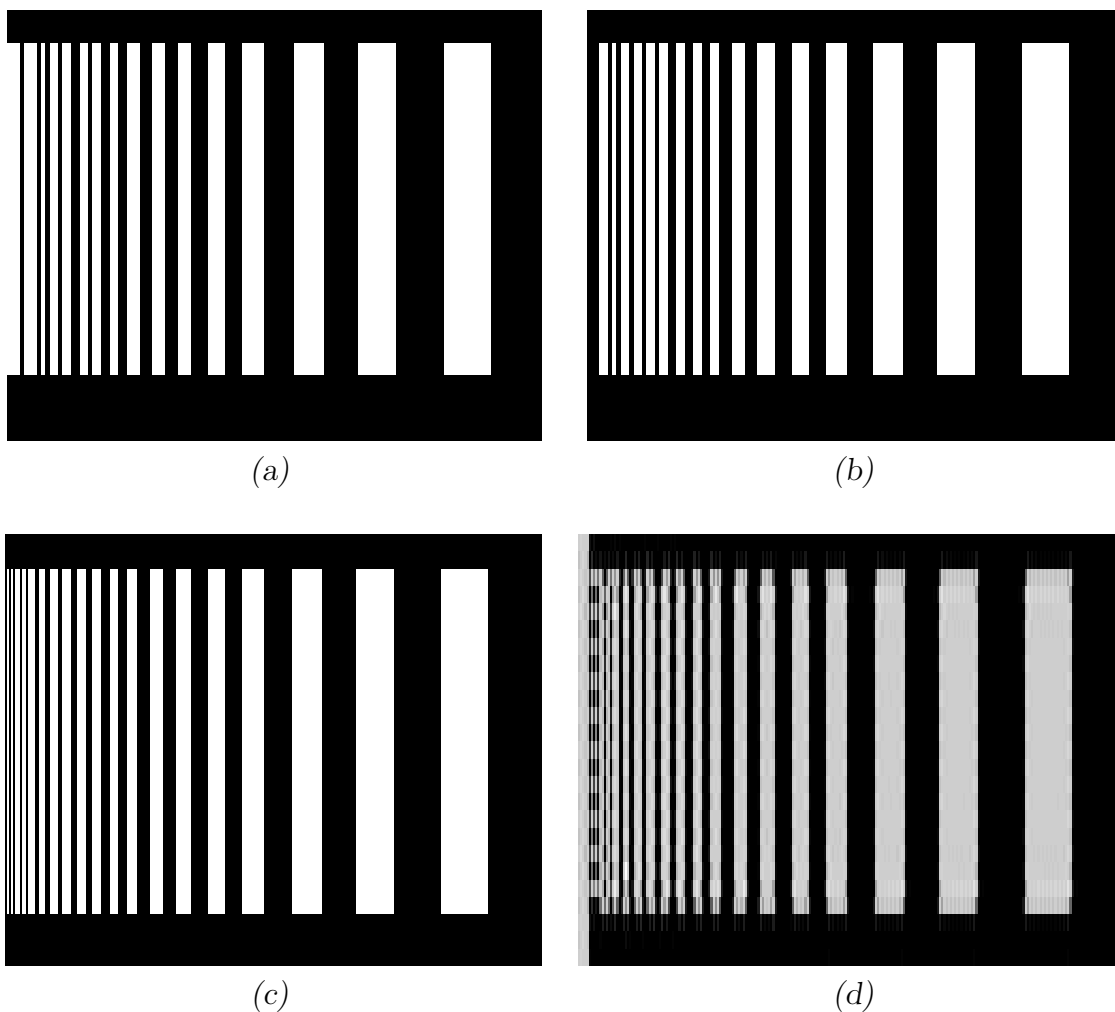


Figure 4.5: Rectangles Comparison. (a) 1st Frame. (b) 2nd Frame. (c) Original Figure. (d) Super-Resolution Image. The size and spacing of the rectangles are drastically distorted in each frame. The super-resolution image recreates each individual rectangle, and at the proper spacing. Though the smaller rectangles are not crisp they are still distinguishable from one another

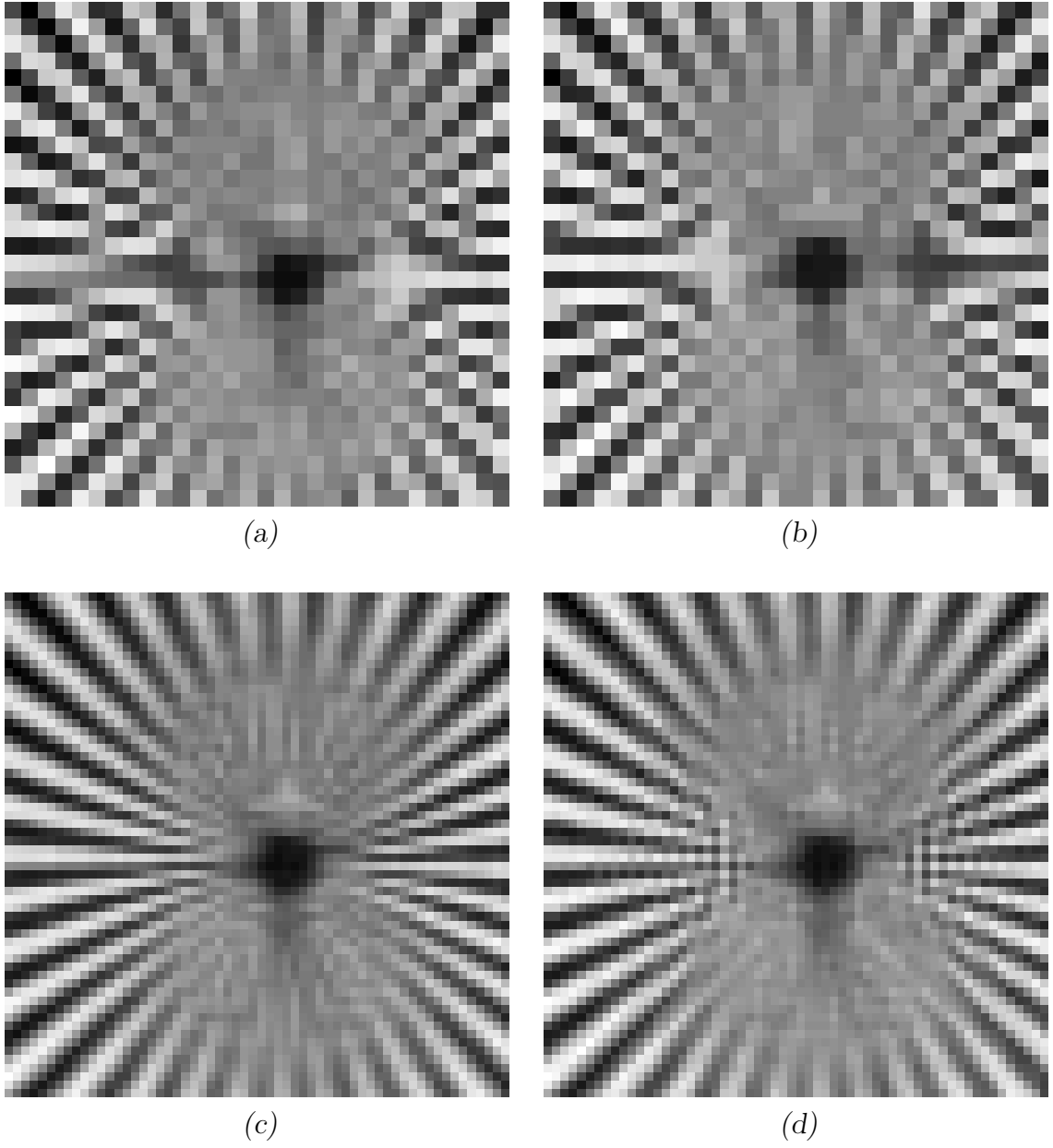


Figure 4.6: Pinwheel Comparison. (a) 1st Frame. (b) 2nd Frame. (c) Original Figure. (d) Super-Resolution Image. The original lines are nearly perfectly reconstructed. The only discrepancies are at the center of the image. Each line is still distinguishable.



(a)



(b)



(c)



(d)

Figure 4.7: Lenna Comparison. (a) 1st Frame. (b) 2nd Frame. (c) Original Figure. (d) Super-Resolution Image. There is a great loss of detail each frame, most noticeable in the feather of the hat, the eyes, and the lips. The detail that is reconstructed, especially in the hat feather is remarkable.



(a)



(b)

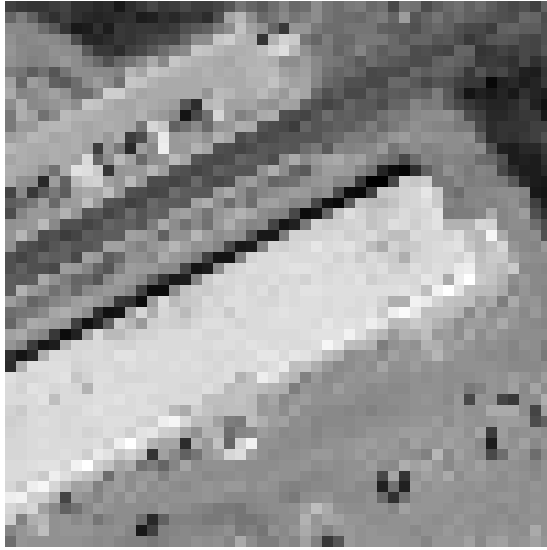


(c)

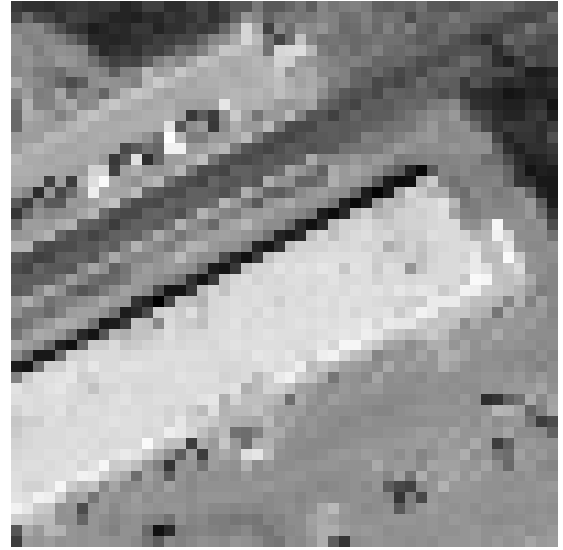


(d)

Figure 4.8: Cameraman Comparison. (a) 1st Frame. (b) 2nd Frame. (c) Original Figure. (d) Super-Resolution Image. The pillars of the building in the background, and the details in the camera are almost perfectly reconstructed. The legs of the tripod have regained their smoothness.



(a)



(b)

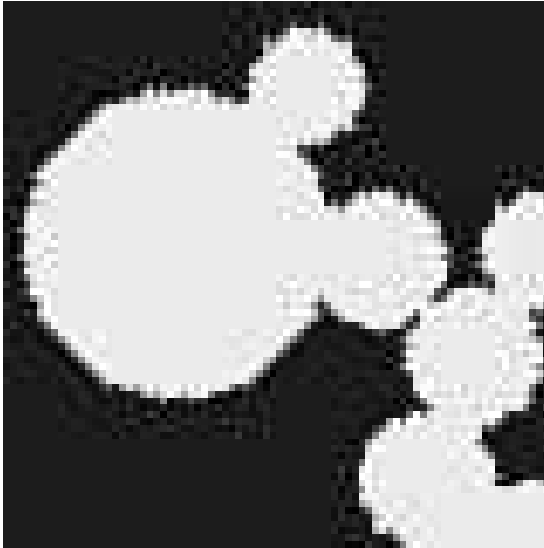


(c)

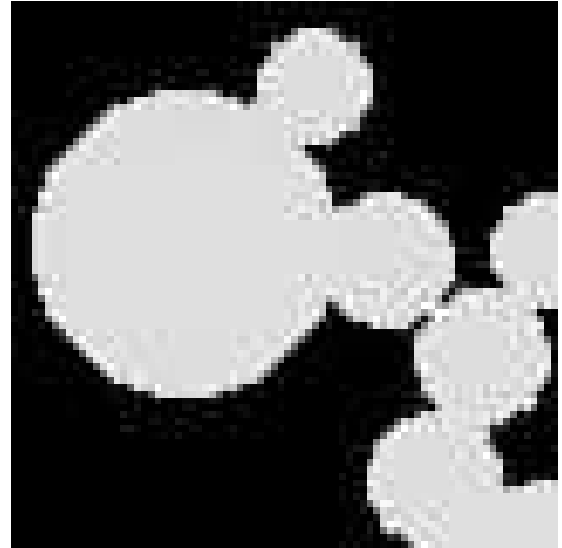


(d)

Figure 4.9: Aerial Comparison. (a) 1st Frame. (b) 2nd Frame. (c)Original Figure. (d) Super-Resolution Image. The building and the road are the only distinguishable features in the frames. In the super-resolution image the parking lot becomes obvious all the way down to the distinction of individual vehicles and parking spaces.

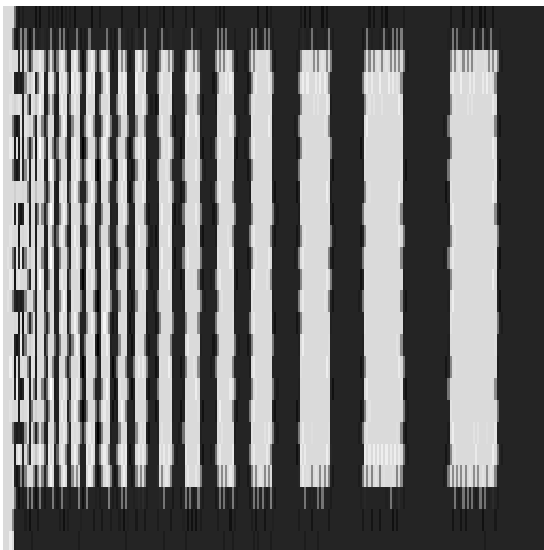


(a)

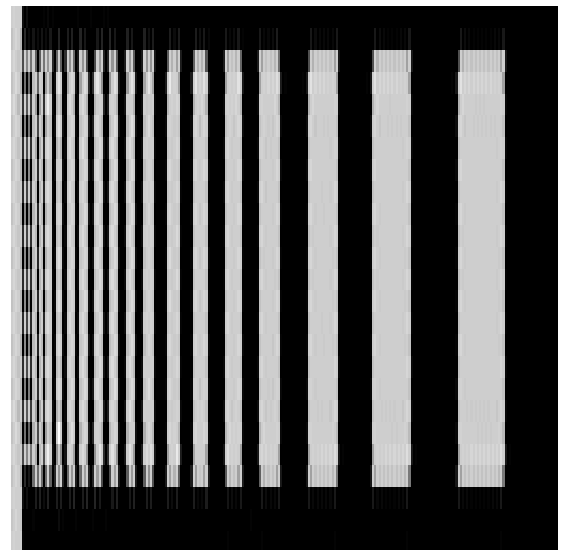


(b)

Figure 4.10: Interpolation Comparison with Circles Image. (a) Cubic Interpolation. (b) Wavelet Interpolation. The borders of the circles are much crisper in the wavelet interpolated image.



(a)



(b)

Figure 4.11: Interpolation Comparison with Rectangle Image. (a) Cubic Interpolation. (b) Wavelet Interpolation. As in the circles, the edges of the rectangles are reconstructed much better in the wavelet interpolation image. The wavelet interpolated image had difficulty exactly recreating the smaller rectangles, but the cubic interpolated image recreates them with severe distortion.



Figure 4.12: Interpolation Comparison with Lenna Image. (a) Cubic Interpolation. (b) Wavelet Interpolation. The cubic and wavelet interpolated images appear similar, but the finer details such as the feather and the lines of hair are clearer in the wavelet image.

most significant in the pinwheel image. The results show that the post-processing step of the algorithm effectively reduces interpolation error.

4.5 Summary

The qualitative analysis of images illustrated that the wavelet based super-resolution algorithm, using the Daubechies 7,9 wavelet system, accurately reconstructed the original image. This reconstruction was significantly better than cubic spline interpolation. The quantitative analysis showed the superiority of the Daubechies 7,9 wavelet system to other wavelet systems and interpolation tech-

	Circles	Rectangles	Pinwheel	Aerial	Lenna	Cameraman
Haar	14.8584	22.2944	890.4712	1.1831e3	1.3695e3	1.8864e3
Daub 9,7	14.6337	23.1060	924.0684	1.1639e3	1.3607e3	1.9166e3
Daub 7,9	14.3272	21.9297	874.0697	1.1264e3	1.3230e3	1.8591e3

Table 4.2: Mean Square Error Comparison of Wavelet Filters.

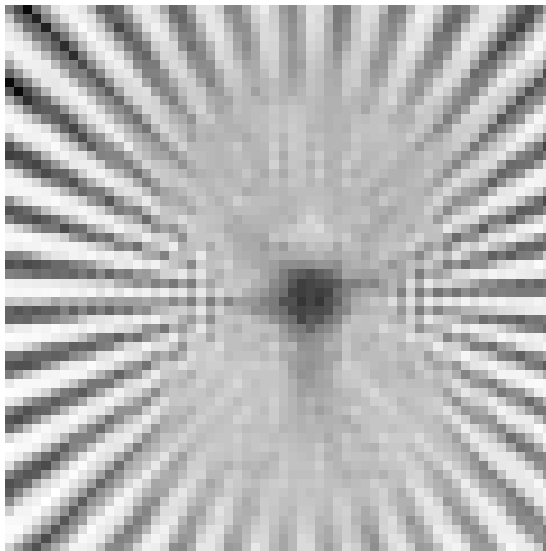


(a)

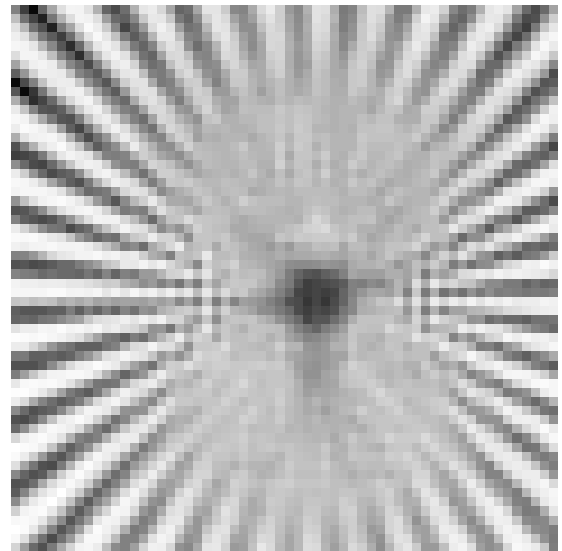


(b)

Figure 4.13: Interpolation Comparison with Cameraman Image. (a) Cubic Interpolation. (b) Wavelet Interpolation. As with Lenna, the cubic interpolated image reproduces general features, but fails to accurately reproduce the finer details. This is most noticeable in the camera detail, along the legs of the tripod, and the shoulders of the cameraman.



(a)



(b)

Figure 4.14: Pinwheel Image Before and After Post-Processing. (a) Before. (b) After. The dark diagonal lines are smoother in the after image

	Circles	Rectangles	Pinwheel	Aerial	Lenna	Cameraman
Daub 7,9	14.3272	21.9297	874.0697	1.1264e3	1.3230e3	1.8591e3
Cubic	18.6622	34.9225	1.3499e3	1.3077e3	1.6213e3	2.3884e3
Linear	19.0132	37.1147	1.8144e3	1.4861e3	1.7778e3	2.5801e3
Nearest	26.4575	50.6360	4.6764e3	2.3484e3	2.3540e3	3.1974e3

Table 4.3: Mean Square Error Comparison of the Daubechies 7,9 Wavelet with Traditional Interpolation Methods.

	Circles	Rectangles	Pinwheel	Aerial	Lenna	Cameraman
Haar	1.0186	1	112.3714	100.2184	89.2220	137.6439
Daub 9,7	1.0586	1	102.3207	94.5569	88.2995	123.4636
Daub 7,9	1.0214	1	95.3491	97.2869	85.7935	126.9404

Table 4.4: L-Infinity Comparison of Wavelet Filters

	Circles	Rectangles	Pinwheel	Aerial	Lenna	Cameraman
Daub 7,9	1.0214	1	95.3491	97.2869	85.7935	126.9404
Cubic	1.0625	1	94.0500	106.4375	116.9357	173.0607
Linear	1	1	91.5000	106.0000	122.5000	192.5000
Nearest	1	1	121.0000	114.6667	148.0000	174.0000

Table 4.5: L-Infinity Comparison of the Daubechies 7,9 Wavelet with Traditional Interpolation Methods.

	Circles	Rectangles	Pinwheel	Aerial	Lenna	Cameraman
Before	14.1116	23.6796	1.1409e3	1.2644e3	1.4101e3	1.9521e3
After	14.3272	21.9297	874.0697	1.1264e3	1.3230e3	1.8591e3

Table 4.6: Mean Square Error Comparison of Post-Processing.

	Circles	Rectangles	Pinwheel	Aerial	Lenna	Cameraman
Before	1.0230	1.0848	160.0123	97.9321	100.7054	132.0056
After	1.0214	1	95.3491	97.2869	85.7935	126.9404

Table 4.7: L-Infinity Comparison of Post-Processing

niques, in terms of mean square error. In terms of L-infinity error the Daubechies 7,9 wavelet was shown to perform better in images with fine detail and Daubechies 9,7 wavelet was shown to perform better in images with coarse information. For the binary images, in terms of L-infinity error, all systems performed about the same. It was also shown that the post-processing of the super-resolution image produced smaller MSE and L-infinity error. The significance of these results is discussed in the following chapter.

V. Discussion and Future Work

5.1 Contributions Of This Thesis

A wavelet-based super-resolution algorithm has been developed that can accurately produce one high-resolution image from two sub-pixel shifted low-resolution frames. The algorithm assumes this shift to be a half-pixel in the x and y direction. The frames are combined into a high-resolution grid according to this shift, and rotated to optimize use of the wavelet transform.

Once in the wavelet domain, missing high-resolution coefficients are calculated using an adaptive interpolation method. The method utilizes the parsimonious clustering property of the wavelet transform to find an optimal set of three or four coefficients. The average of this coefficient set determines the value of the missing coefficient. Once all missing coefficients have been calculated, the new high-resolution image is brought out of the wavelet domain and re-rotated into its original orientation.

Due to interpolation error, certain edges in the image are not properly reconstructed. A post-processing step uses the wavelet transform to slightly subdue horizontal and vertical edges, and enhance diagonal ones. The final high-resolution image shows accurate reconstruction of most high-resolution details.

Different wavelet systems were applied to this algorithm to determine which wavelet system most accurately reconstructed the the high-resolution image. The MSE and L-infinity errors of each of these wavelets were used to quantitatively determine the accuracy of reconstruction. The Daubechies 7,9 wavelet system was shown to perform the best in terms of mean square error. In terms of L-infinity error the Daubechies 7,9 wavelet system performed better on images with fine detail, while the Daubechies 9,7 wavelet system performed better on images with coarse information. Using this wavelet system, the algorithm was then compared to other traditional interpolation techniques. Chapter 4 illustrated the superior performance

of the wavelet-based super-resolution algorithm to these techniques, both visually and in terms of MSE and L-infinity error.

5.2 Potential For Future Research

This thesis developed an accurate wavelet-based super-resolution algorithm that operates on two frames at one specific sub-pixel shift. A more robust algorithm could be developed to handle multiple frames at any arbitrary sub-pixel shift.

Since many good image registration and compression algorithms use wavelets, the super-resolution algorithm developed here could be combined with either, or both, of these, to provide a more computationally efficient overall system. This algorithm can, and should be, extended into other more robust applications.

Application of the RDWT was only done at the first scale. More accurate results may be produced by working at higher scales. The application of the non-redundant DWT may also effectively achieve super-resolution.

The algorithm was only designed to restore resolution lost by the finite CCD array. It may be extended to restore resolution lost by the rest of the camera's optics. By accounting for all lost resolution, and incorporating an accurate image registration algorithm, a system could be developed to apply directly to a real video sequence.

Appendix A. Determination of Missing Pixel Locations

Missing pixel locations are determined by convolving a mask with known coefficient locations, shown in Figure 1.1(a). A temporary copy is made of the wavelet domain image. All coefficients in this image are set to a value of 1. Convolution of this image of 1's with the mask will produce numbers ranging from 0-4 in the locations of the missing coefficients. If the result is a 0 or a 1, it is assumed that no pixel coefficient exists in this location. Coefficients with a value of 1 are then set to 0. This accounts for erroneous coefficients produced along the edges of the rotated image. A coefficient is assumed to exist in locations that have a value of 2 or greater. These are all set to a value of one. The indices of all locations with a value of one are recorded. The adaptive interpolation scheme of Section 3.2.2 picks the 4 coefficients surrounding each recorded pixel location to determine the missing coefficient value. Figures 1.1(b) and 1.1(c) show the before and after result of the interpolation. New coefficients are represented by *'s and original coefficients are represented by diamonds.

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

(a)

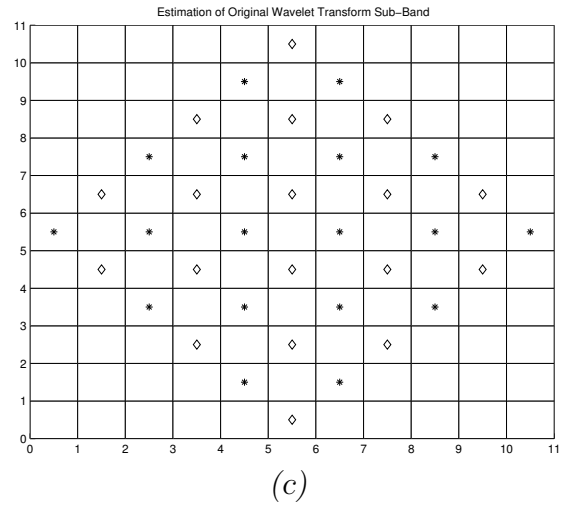
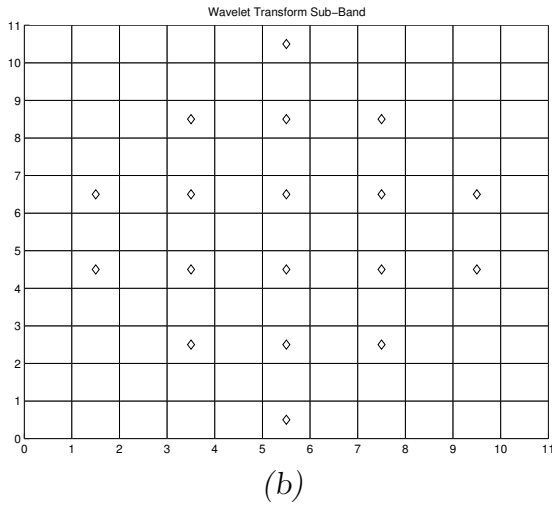


Figure 1.1: Convolution of Mask with Sub-Band. (a) Mask. (b) Wavelet Transform Sub-Band of Test Frames. (c) Wavelet Transform Sub-Band with Missing Values Added.

Bibliography

1. Baker, S. and T. Kanade. *Limits on Super-Resolution and How to Break Them*. Technical Report, Pittsburgh, PA: The Robotics Institute, Carnegie Mellon University, 2002.
2. Baker, Simon and T. Kanade. *Super Resolution Optical Flow*. Technical Report CMU-RI-TR-99-36, Pittsburgh, PA: Robotics Institute, Carnegie Mellon University, October 1999.
3. Brown, R. K. *Image Registration using Redundant Wavelet Transforms*. MS thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, March 2001.
4. Burrus, C. S., et al. *Introduction to Wavelets and Wavelet Transforms: A Primer*. Upper Saddle River, NJ: Prentice-Hall, 1998.
5. Chen, Y.W. and N. Mendoza. "A Heuristic Approach to Image Superresolution," *Proceedings of 2nd IEEE Workshop on Intelligent Signal Processing 2001*, pressed
6. Daubechies, I. "Orthonormal Bases of Compactly Supported Wavelets," *Communications on Pure and Applied Mathematics*, 41:909–996 (July 1988).
7. di Francia, G. Toraldo. "Resolving Power and Information," *J. Opt. Soc. Am.*, 45:497–501 (July 1955).
8. Elad, M. and A Feuer. *Super-Resolution Reconstruction of Continuous Image Sequences*. Technical Report 1042, Haifa 32000, Israel: The Technion Israel Institute of Technology, July 1996.
9. Elad, M. and A. Feuer. "Restoration of a Single Superresolution Image from Several Blurred, Noisy and Undersampled Measured Images," *IEEE Transaction on Image Processing*, 6(12):1646–1658 (December 1997).
10. Huang, T.S. and R. Tsai. "Multi-Frame Image Restoration and Registration," *Advances in Computer Vision and Image Processing*, 1:317–339 (1984).
11. Hunt, B.R. "Super-Resolution of Images: Algorithms, Principles, Performance," *Int. J. Imaging Sys. and Tech.*, 6:297–304 (Winter 1995).
12. Keren, D., et al. "Image Sequence Enhancement Using Sub-Pixel Displacements." *Proceedings of the 1988 Conference on Computer Vision and Pattern Recognition*. 742–746. 1988.
13. Lim, J. S. *Two-Dimensional Signal and Image Processing*. New Jersey: Prentice Hall, 1990.
14. Mallat, S. and W. Hwang. "Singularity Detection and Processing with Wavelets," *IEEE Transactions on Information Theory*, 38(2):617–643 (1992).

15. Mallat, S. and S. Zhong. "Characterization of Signals from Multiscale Edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:710–732 (July 1992).
16. Mallat, S. G. "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693 (1989).
17. Manfra, J. L. *Translation and Rotation Invariant Multiscale Image Registration*. MS thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, March 2002.
18. Mendoza, N., et al. "A Hybrid EA Approach to Multisensor Image Superresolution," *Tech. Bul. Of University Ryukyus*, 69–77 (2001).
19. Mendoza, N., et al. *A Systematic SC-Based Methodology for Super-resolution of Video Sequences*. Technical Report 143, Okinawa-ken, 903-0129, Japan: University of the Ryukyus, May 2001.
20. Mendoza, N.E., et al. "A real multi-parent tri-hybrid evolutionary optimization method with applications in the resolution of overlapping signals," *Proceedings IEEE Int. Conf. Industrial Control and Instrumentation*, 2837–2842 (2000).
21. Peleg, S., et al. "Improving Image Resolution Using Subpixel Motion," *Pattern Recognition Letters*, 223–226 (1987).
22. Schultz, R. and R. Stevenson. "Extraction of High-Resolution Frames From Video Sequences," *IEEE Transactions on Image Processing*, 5(6):996–1011 (1996).
23. Sementilli, J. J. *Suppression of Artifacts In Super-Resolved Images*. PhD dissertation, University of Arizona, 1993.
24. Strang, G. and T. Nguyen. *Wavelets and Filter Banks*. Wellesley, Cambridge, 1996.
25. Ur, H. and D. Gross. "Improved Resolution From Subpixel Shifted Pictures," *Computer Vision, Graphics, and Image Processing*, 54(2):181–1863 (1992).

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY) 25-03-2003		2. REPORT TYPE Master's Thesis		3. DATES COVERED (From - To) Aug 2001 - Mar 2003	
4. TITLE AND SUBTITLE REDUNDANT DISCRETE WAVELET TRANSFORM BASED SUPER-RESOLUTION USING SUB-PIXEL IMAGE REGISTRATION				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Ward, Daniel L, Second Lieutenant, USAF				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way, Building 640 WPAFB OH 45433-7765				8. PERFORMING ORGANIZATION REPORT NUMBER AFIT/GE/ENG/03-18	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) USAF C2B Concepts and Research Division Attn: Lt Col Stephen Matechik 238 Hartson Street Hurlburt Field, FL 32544-5237 MAJCOM: AFSOC DSN: 579-8239 e-mail: Stephen.Matechik@hurlburt.af.mil				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The limited resolution of video imagery taken by aircraft, over geographical areas of interest, hinders the accurate extraction of useful information. The frame resolution of the video is determined by the camera that created it. Information exists about the camera which can be used to increase frame resolution beyond the resolution capability of the camera. This is achieved by a process called {it super-resolution}, which uses multiple low-resolution video frames to create one high-resolution image. Paramount to this super-resolution process is the alignment of the low-resolution frames. Frame alignment is achieved by image registration. Many good image registration algorithms utilize a wavelet domain analysis of the images. It is theorized that an accurate super-resolution algorithm can be created that also utilizes wavelet domain analysis. By achieving both image registration and super-resolution in the wavelet domain, a more computationally efficient overall system may be developed. This thesis explores the possibility of wavelet-based super-resolution. An algorithm is developed that restores resolution lost by the CCD array of the camera. Two low-resolution frames, registered at a specific shift from one another, are used to create one high-resolution image. The algorithm is compared to other traditional interpolation techniques, and different wavelets systems are applied to determine which one works best.					
15. SUBJECT TERMS Wavelet Transform, Resolution, High Resolution, Low Resolution, Video Frames, Aerial Imagery, Image Registration, Image Processing					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Roger L. Claypoole Jr., Maj, USAF (ENG)
U	U	U	UU	72	19b. TELEPHONE NUMBER (Include area code) (937) 255-6565, ext 4620; e-mail: Roger.Claypoole@afit.edu